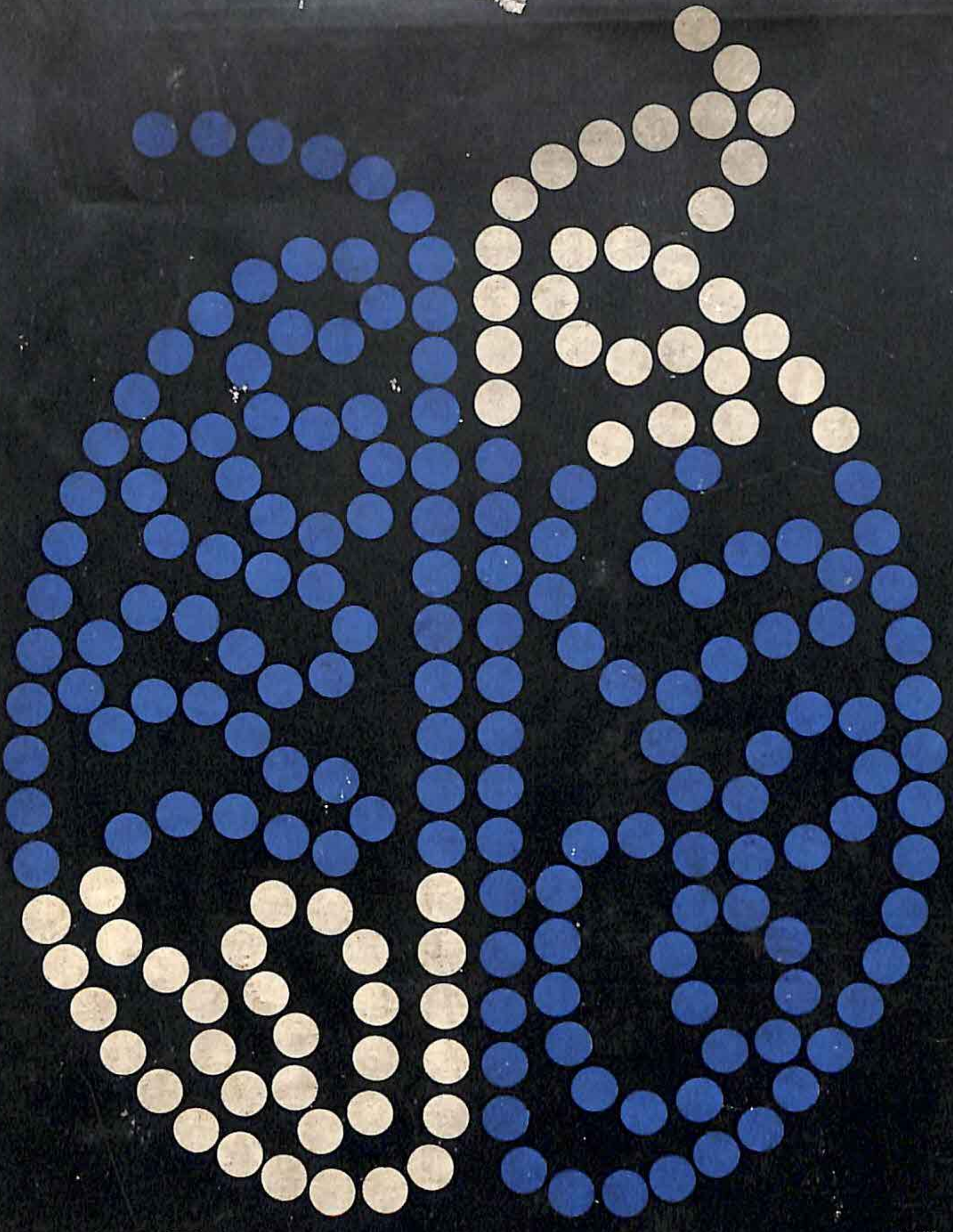

William K. Estes

learning theory and mental development

Learning Theory and Mental Development

c

Learning Theory and Mental Development

WILLIAM K. ESTES

The Rockefeller University
New York, New York

Calcutta

1970

ACADEMIC PRESS *New York and London*

15-1-72
२२९९

COPYRIGHT © 1970, BY ACADEMIC PRESS, INC.
ALL RIGHTS RESERVED
NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, RETRIEVAL SYSTEM, OR ANY
OTHER MEANS, WITHOUT WRITTEN PERMISSION FROM
THE PUBLISHERS.

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

1544
EOT

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
Berkeley Square House, London W1X 6BA

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 74-117636

PRINTED IN THE UNITED STATES OF AMERICA

C.E.R.T., West Bengal
Date 15-1-72
Acc. No. 2299

Contents

Foreword ix

Introduction 1

PART A. THEORETICAL BACKGROUND

I. Modern Learning Theory 7

1. The Elements of Association 8
2. Conditions of Association 9
3. Conditions of Transfer to Long-Term Memory 9
4. Retrievability from Long-term Memory 10
5. Motivation and Response Evocation 10
6. Conditions of Retention and Transfer 12
7. Probability Learning 13
8. Formation of Hierarchies of Stimulus and Response Units 14
9. Learning to Learn 15

II. Interpretations of Standard Learning Situations 17

1. Classical Conditioning 17
2. Operant Conditioning 20
3. Discrimination Learning 23
4. Learning Set 26
5. Paired-Associate Learning 27

III. Individual Differences and Developmental Processes 30

1. The Concept of Learning Ability 30
2. Development of Learning Abilities 32

PART B. EMPIRICALLY ORIENTED RESEARCH

IV. Heritability and Development of Learning Ability 45

1. Genetic Studies 45
2. Developmental Studies 46

V. Empirical Comparisons between Normal and Retarded Populations 52

1. Methodology 52
2. Comparative Experimental Studies 58

VI. Modifying Rates of Learning and Forgetting 71

1. Classical and Avoidance Conditioning 71
2. Operant Conditioning 77
3. Discrimination Learning 78
4. Paired-Associate Learning 84
5. Educational Situations 86

PART C. APPLICATIONS OF LEARNING THEORIES

VII. Hull's Theory 97

1. Major Concepts and Postulates 97
2. Interpretation of Individual Differences in Learning Ability 100
3. Direct Applications of the Hull-Spence Theory to Research on Learning Ability 104

VIII. Ellis's Stimulus-Trace Theory 111

1. Background and Assumptions of the 1963 Formulation 111
2. Tests and Implications of the Trace Theory 113
3. Reformulation in Terms of Memory Processes 115
4. Research Deriving from the Revised Theory 116
5. Comment 119

IX. Denny's Elicitation Theory 121

1. Elicitation of the Correct Response 122
2. Immediate Knowledge of Results 123
3. Stimulus Generalization and Positive Transfer 124
4. Motivation (Elicitation Hierarchy) 124
5. Sequential Building 124

X. The Social Learning Approach 126

1. Social Learning Theory 126
2. Research in "Social Learning" 127
3. A Tentative Formalization 130
4. Comment 130

XI. Field Theory 132

XII. The Operant Conditioning Approach 137

1. Concepts and Principles 137
2. Approach to Mental Retardation 139

XIII. Contemporary Discrimination Learning Theories 146

1. Background and Experimental Context of the Attentional Theory of Zeaman and House 146
2. A Two-Stage, Attentional Model for Discrimination Learning 151
3. Attention Theory Applied to Mental Retardation 166
4. Discrimination Shift Studies 169
5. General Comments on Discrimination Theories and Related Experimental Analyses 180

XIV. General Comments on Methodology 183

References 189

Author Index 207

Subject Index 213

Foreword

The preparation of this volume was supported in part by a grant from the Joseph P. Kennedy, Jr. Foundation. The idea of conducting a critical analysis of research on mental retardation in relation to theories of learning was proposed to me originally by Nicholas Hobbs on behalf of the Foundation. I was by no means entirely satisfied with my qualifications for this assignment, having worked exclusively on one side of the relationship. However, the task offered an interesting challenge and it seemed to me that it would be worth trying to appraise what has been done and what might be done to exploit learning theory in this research area. In further discussions with Dr. Hobbs and communications with Mr. Sargent Shriver, then Director of the Foundation, I agreed to undertake a critical analysis and to attempt to complete it within a year. The year in question, which was to have been 1966, stretched eventually to three, primarily because none of us appreciated the extent to which the literature in this area is burgeoning. Further, it became increasingly evident to me as the work proceeded that research on mental retardation cannot meaningfully be separated from research on the development and organization of learning abilities. The title of this monograph reflects the enforced shift of emphasis.

This volume is intended to make the results of the study available to investigators of mental development, to students preparing for research in this field, and to readers with general interest in theories of learning and their applications.

I hope that the emphasis on criticism in this report does not outweigh in the minds of readers the substantial appreciation due the investigators whose imaginatively conceived and energetically prosecuted research programs are producing substantial progress toward the building of a new discipline. I thought that it was my task, not to mete out praise or blame, but to concentrate on the possibilities of sharpening and advancing theoretical analysis of the development and maldevelopment of learning abilities. It should be kept in mind that only work which has been considered to be of some merit is cited or discussed in this report. A total of 1770 books and articles were examined closely enough to be entered in our card file. Those considered to add little information and those valuable for other purposes but not germane to the present enterprise are not cited in this volume.

The relatively extensive survey of the literature was possible only with the assistance of Katherine W. Estes, who also smoothed the final copy considerably with her editorial pencil. Mariana Fry Weidner also contributed to the survey in the earlier stages. The transcription of my notes and reflections from many yards of often all but indecipherable dictaphone tape was begun by Winifred Woltering and completed by Edith Skaar. Miss Skaar also deserves credit for careful typing of the final manuscript.

Learning Theory and Mental Development

Introduction

The word "and" in the title of this volume signifies a logical product, not a logical sum. Merely to survey contemporary learning theory requires a substantial book in itself (Hilgard & Bower, 1966). Individual presentations of the major contemporary learning theories could easily fill out a full shelf; and at least the remainder of a sizable bookcase would be required to accommodate the accumulated empirical information and related theory concerning the development of intelligence. But the interaction of learning theory with the study of mental development has been so slight that an extensive search of the literature yields material for only a slender volume.

This state of affairs is rather striking in view of the common tendency virtually to equate intelligence with learning ability. Note, however, that this identification has never been documented by research. The principal trends in both the study of mental organization and the study of learning have been away from conceptions of either learning ability or intelligence as unitary entities. Further, Guilford (1967) notes that, not only have his extensive investigations yielded no factor interpretable as learning ability but, "In fact, one could look a long way in the literature and not find such a factor" [p. 268].

Yet it would be absurd to conclude that learning and intelligence are unrelated. One can measure an individual's intelligence at any given age only by sampling the skills and knowledge he has acquired. Differences in performance between individuals following similar opportunities to learn are, by definition, our only source of information regarding learning ability. They are also our principal basis for measuring intelligence. In both cases, one recognizes that test performance must depend on genetic and maturational factors as well as on experience, but scarcely a start has been made toward separating these contributions.

The reaction to this frustrating situation on the part of a few investigators, especially in the area of language development, has been to dally with the possibility that the development of language, and therefore the enormous range of intellectual performance dependent on language, is independent, if not of learning in a broad sense, at least of the kinds of learning processes studied in the laboratory (Lenneberg, 1967).

This view would seem to be tenable only so long as an investigator limits himself to the study of the normal course of development. Within this limitation, purely developmental theories have their attractions, perhaps the most outstanding example being Piaget's theory of intellectual development (1926, 1952, 1967). However, a major part of the research on the development of intellect is instigated by problems arising when development deviates from the "normal." It is of no help in dealing with, for example, the mentally retarded to be provided only a characterization of these individuals in terms of the stage of mental development they have reached. It is essential for social as well as scientific reasons to exploit every possibility of preventing or remedying mental deficiency. Substantial progress in these respects can be hoped for only to the extent that individual differences in intellectual performance can be traced to specific causes. Thus it is only natural that psychological research on mental retardation has concentrated more and more heavily upon the investigation of learning processes.

The rather sudden acceleration of this type of research has been initiated, not by broad strategic considerations, but by a combination of increasing pressure of practical problems and the appearance of major financial support by the National Institutes of Health, the Kennedy Foundation, and other agencies, following the appearance of the report of the President's Panel on Mental Retardation (1962). Not surprisingly research has advanced on a jagged front, much being purely empirical in conception, some directed by various extant theories.

The large volume of research on learning currently reported in such publications as *Advances in Child Development and Behavior* (Lipsitt & Spiker, 1963, 1965, 1967) and the *International Review of Research on Mental Retardation* (Ellis, 1966a, b, 1968) is surely useful. But it is important to ask whether these increasingly substantial investments of research effort could be made still more effective. Given that support is substantial, the main possibility of increasing effectiveness must lie either in methodology or in the relation between research and theory. The former has been receiving increasing attention during the past few years, starting with reviews by Denny and others presented in Stevens and Heber (1964), and followed by various critical articles in the *Advances* and the *International Review*.

It is the main purpose of the present study to give similar critical attention to the relationship between learning theory and research concerned with mental development—both to normal development, with special attention to individual differences in learning abilities, and to the types of deviations from the norm which define the area of mental retardation.

In order to make clear the basis for my interpretation and judgments, I have first assembled a summary of the main concepts and principles of contemporary learning theory as I see them, at the same time trying to indicate what resources modern learning theory may have to offer to research on mental development. Then I have proceeded to review principal lines of research on the development and maldevelopment of learning abilities, and to interpret the research methods and results in terms of this body of theory.

The third main section of this report concerns the extent to which research on the development of learning abilities has been organized and directed by theory. In the course of attempting to deal with this question it was necessary to examine the particular theories which have been involved to appreciable extents and to seek to determine how useful they have been in practice. In concluding the final section, I have abstracted some implications of the study regarding research approaches that might more fully utilize available theoretical resources.

Any analysis of the literature bearing on individual differences in learning ability must distinguish two main types of relationship to learning theory. The first has to do with the way in which individual differences are represented within any given learning theory, the second, with the theoretical interpretation of individual differences in various types of intellectual performance. Literature pertaining to the former aspect has been reviewed by Glaser (1967). This literature is far from extensive, and as a consequence learning theorists have frequently been sharply criticized for giving little explicit attention to the representation of individual differences within models for learning (e.g., Hirsch, 1962).

It is certainly true that learning theorists are most often concerned with relationships which are believed to be common to all members of a species and that in formalizing their theories they nearly always present theoretical functions which are assumed to hold for homogeneous populations of organisms. However, it is easy to misunderstand the sense in which homogeneity is assumed. Most often homogeneity is merely a simplifying assumption, introduced in order to permit the beginning of mathematical analysis of what, without rather severe simplification, would be an unmanageable tangle of relationships (see e.g., Estes, 1964, pp. 91-93). It is frequently, if not always, clearly recognized by learning theorists that this phase must be followed later by the relaxation of simplifying assumptions and examination of the consequences for theoretical predictions which arise from individual differences in learning parameters (see, e.g., Trabasso & Bower, 1968; Zeaman & Kaufman, 1955).

To my knowledge the first interpretation of mental development in terms of learning theory was undertaken by Thorndike (1931). Thorndike proposed to account for the progressive increase in intellectual capacity from the lower organisms to man in terms of the quantitative increase in the number of available associative connections. Presumably individual differences in learning and intellectual ability among human beings could also be attributed to variation in number of interneuronal connections. According to Thorndike the quantitative increase in the number of available neural connections permits the appearance of ideas, insight, and reasoning. But effective intelligence is a function of the number of associations actually formed. Since one would expect a direct relationship between the amount of knowledge an individual has acquired and the number of associative connections formed in his nervous system, Thorndike believed his analysis to be supported by the well-known positive correlation between measures of information or vocabulary and other measures of mental ability.

Over the 30-40 years since Thorndike's pioneering efforts, there has been considerable change in the frame of reference within which learning theorists approach problems of individual differences, but very little advance in the analysis of specific problems. In his general comments on a symposium concerned with individual differences in learning, Melton reported (1967) as the consensus of the symposium contributions that research on individual differences in learning must be guided by theories of human learning and performance, and that progress in this area must depend upon the framing of hypotheses regarding individual difference variables "*in terms of the process constructs of contemporary theories of learning and performance* [p. 239]."[1] Analysis of the extent to which this prescription is actually being effectuated in contemporary research, or could be in the near future, is a principal concern of the present volume.

---

[1] A similar view has been expressed by Scott (in press), with documentation in terms of a careful review of relevant literature.

Part A

Theoretical Background

I. *Modern Learning Theory*

Borrowing an expression from Miller (1959), we might characterize the viewpoint of the present review as that of a liberalized association theory. The history of learning theory, in this country at least, has been one of continuing elaboration of association theory, with major contributions coming during each succeeding generation from more transitorily influential developments, for example, gestalt psychology, information theory, and, most recently, computer analogies. Although association theory has its limitations, and other views their points of special appeal, it is within the associationistic framework that the preponderance of research has been conducted leading to detailed interpretations of numerous varieties of learning.

It should be recognized that the conception of the organism in modern association theory is considerably richer than a bundle of reflexes or stimulus-response bonds. The specific version which, I believe, now marshals broadest support is a stimulus-response formulation at the level of performance, but not at the level of learning, conceived as an inference from performance. In order to be rigorously testable and to be practically useful in the guidance and management of training and education, learning theory must provide detailed analyses and predictions of the changes in an individual's dispositions to respond to stimuli or classes of stimuli that develop as a function of experience. Thus, the formulation I shall outline begins and ends the treatment of each type of learning situation with analyses of performance in stimulus-response terms, following the general approach of Guthrie (1935), Hull (1943), and Skinner (1938). However, the associative process is conceived to be, not a matter of switching connections of a response from one stimulus to another, but rather the learning of relations between events. The connection between learning and performance involves feedback relationships in which the process of response selection is continually modified by anticipations of rewarding and punishing consequences.

It will become clear, I trust, in the following presentation that only a drastic misconception of modern association theory could put it in opposition to views which emphasize the importance of structure and organization in behavior as, for example, the "structure-of-intellect" theory of Guilford (1967). On the contrary, the basic properties of the associative process lead to the formation of progressively higher-order perceptual and motor units. The result of an extended period of learning is not simply a proliferation of associative relations between elementary stimuli and motor reactions, but rather a hierarchical organization of associative relations between stimulus and response patterns of varying levels of complexity.

In this section I shall sketch the set of concepts and assumptions drawn from contemporary theories of learning which I shall use as a basis for interpreting research on learning in relation to mental development. While I assuredly would not claim a concensus among learning theorists on the framework presented, I believe the principal assumptions have broad empirical support and that, subject to differences in emphasis and particular terminology, they constitute the framework within which much current research on learning is being conducted. For purposes of exposition I shall first list the principal assumptions with only brief elucidation, then give illustrative interpretations of a few standard types of learning experiments, and finally consider applications to problems of mental retardation and acceleration.

1. the elements of association

I take the term *learning* to refer always to a change of state of the organism inferred from performance—more specifically, to any relatively enduring change in behavioral dispositions which occurs as a result of previous experience. At a conceptual level, learning is represented by the formation of associations between representations of behavioral events which have occurred in temporal contiguity. By *behavioral event* I denote the organism's reaction to any stimulus input, including both perceptual and motor aspects.

This formulation represents a change from my earlier theoretical papers (Estes, 1950, 1959) which were developed within the stimulus-response framework of Guthrie (1935) and Hull (1943), with the basic assumption that learning proceeds by stimulus substitution in S-R units. Now, I am persuaded that a wide variety of phenomena which had received little investigation at the time of those earlier formulations can be most simply interpreted in terms of a process more closely akin to the conception of S-S conditioning (Spence, 1951). Examples are conditioned anticipation of aversive events and of reward (Estes & Skinner, 1941; Estes, 1948), latent learning (MacCorquodale & Meehl, 1954), incidental learning (Postman, 1964), and conditioning under curare (Solomon & Turner, 1962).

The associations between representations of behavioral events are assumed to be characterized by multiplicity and random fluctuations in state of activity in accord with the mechanisms developed within earlier expositions of stimulus sampling theory (see Estes, 1959; Neimark & Estes, 1967). Thus, of the total number of elementary associative connections which might be established between representations of two behavioral events, only a subset may be established on any single learning trial, and of these only a random sample will be in an active state on any single later trial.

2. conditions of association

The contiguous occurrence of two behavioral events in temporal succession is assumed to be the necessary and sufficient condition for establishing temporary associations, or, in other terms, for entering the relationship between these events in short-term memory. Following a contiguous occurrence, the associations revert to their initial state at a random rate over time, generating a short-term retention curve. During the short-term retention interval, which is normally a matter of seconds, some of the associations may become consolidated instead of reverting to the initial state. In this case the associations will be reactivated on a later occurrence of one of the associated events and we say that there has been a transfer of information to long-term memory storage. If this transfer does not occur, then the associations will not be reactivated by later occurrence of either event and we say that the information has been lost from memory.

3. conditions of transfer to long-term memory

Perhaps the most important condition for consolidation is an interval free of interference from the occurrence of similar events (that is, events which would involve associative elements overlapping with those just established) or unusual disturbances of the nervous system such as those produced by electroconvulsive shock. Further, transfer to long-term memory is facilitated by overt or covert repetition of the associated behavioral events, that is, the type of activity termed rehearsal. Finally, there is considerable evidence to indicate that the activity which facilitates transition to long-term memory is initiated only if the second of two associated events resolves some uncertainty on the part of the learner (see, e.g., Egger & Miller, 1962; Trabasso & Bower, 1968). Suppose, for example, that a stimulus compound consisting of components A and B is presented to an organism, followed by event C. If as a result of previous learning, A has been established as a reliable predictor of C, then associations between B and C will not become consolidated in long-term memory.

4. retrievability from long-term memory

Given that associations have been formed and consolidated between two behavioral events, A and B, which have occurred contiguously, a later occurrence of the sequence AB gives rise to the reaction we term recognition. Whether a later occurrence of A alone will activate a representation of B in memory depends upon a number of conditions, including number of previous repetitions, duration of the retention interval, and activities on the part of the learner which modify the form of memory storage. The principal role of the retention interval in relation to retention loss may be to permit the formation of interfering associations between A and other events. The principal function of repetition is to permit an increase in the number of associations as a result of the random-sampling process. The ability of an individual upon the later occurrence of A to reinstate the overt response components of event B may, at least in the case of human beings, be greatly facilitated by a type of activity termed recoding (Bower, 1967; Greeno, 1969; Lawrence, 1963). Of special importance is the conversion of visual to auditory representations by means of verbal rehearsal. For most individuals, at least, the visual components of memory of a previous experience contain relatively little information. Consequently a great gain in accuracy of recall is possible if, during the short-term memory interval, various properties or aspects of what has been seen are verbalized and converted to auditory storage.

5. motivation and response evocation

It is assumed that, except perhaps for the simplest reflexes, response evocation depends upon the joint action of stimulus input from receptors and input from drive mechanisms. The latter presumably involve the midbrain centers which are known to exert rewarding or punishing effects upon direct electrical stimulation (see, e.g., Miller, 1963; Olds, 1961), but we shall not attempt any detailed physiological interpretation here. Originally, these specific mechanisms are activated by unconditioned stimuli, for example, the mechanism associated with hunger is activated by the taste of food, the mechanism associated with pain by impinging traumatic stimulation. The result of activity of a drive mechanism is to generate what may be termed facilitatory or inhibitory feedback. For example, the taste of food initiates facilitatory feedback which then summates with the inputs of stimuli associated with ingestive behaviors and leads to the continuation of consummatory activity. If, while the animal is eating, it receives a painful shock, the result is initiation of inhibitory feedback from the negative drive system. An important aspect of this feedback is that it reciprocally inhibits the activity of the currently active positive drive system, thus cutting off facilitatory feedback and leading to the cessation of consummatory behavior. At the same time, the feedback from the negative drive center summates with unconditioned stimuli for such reactions as attack or flight.

By associative learning, proceeding in accord with the assumptions of Section 1 above, control of these positive and negative drive mechanisms is extended to stimuli which have preceded the original unconditioned stimuli. If the sight of food has preceded the taste on an earlier occasion, then the sight of food will lead to reinstatement in memory of a representation of the unconditioned stimulus, which will in turn initiate facilitatory feedback, thus increasing the likelihood that the response sequence leading from the sight of food to ingestion will again run off without interruption. On the other hand, upon later occurrence, a stimulus which has preceded electric shock will activate a central representation of the shock experience, thus leading to negative drive activity and consequently to reciprocal inhibition of positive drive systems which have been supporting ongoing positively motivated behavior. This suppression of ongoing behaviors by a stimulus which has preceded a traumatic event is the much investigated conditioned emotional response (CER) first analyzed by Estes and Skinner (1941).

The result of this combination of motivational and associational mechanisms is that after some learning experience the organism's behavior is continually modulated by anticipations of rewards or punishments, behavior sequences leading to increases in positive feedback being preferentially selected and behavior sequences leading to decreases in positive feedback or to punishment being inhibited. A consequence of these mechanisms which is of special importance for discrimination learning is that in any situation involving differential contingencies between some stimuli or stimulus aspects and reward, and between other stimuli or stimulus aspects and nonreward or punishment, the organism will learn to approach the former and to avoid the latter.

6. conditions of retention and transfer

The three principal conditions of retention loss that we can identify are (*a*) spontaneous, random fluctuations in the state of activity of elementary associations, (*b*) changes in stimulus context, and (*c*) interfering learning.

The first of these refers to the assumptions (see Section 1 above) that the associations established upon contiguous occurrence of any two behavioral events revert to their initial state at a random rate over time unless consolidation occurs. So far as is known this short-term memory loss is largely an autonomous process that can be modified only by varying the number of associations established upon a given learning experience. If, of three individuals exposed visually to the same paired-associate item, one simply looks at each member, one also verbalizes each member aloud, and one in addition conjures up a visual image of each stimulus in some context, we should expect the third individual to establish the largest number of temporary associations and the first individual the least; thus they would rank in this order with respect to recognition or recall after a short retention interval.

The second factor refers to the fact that situations are never reproduced identically and that, other things equal, the longer the delay between two successive presentations of a particular stimulus or stimulus complex to an organism, the more the context is likely to have changed. If two stimuli have occurred in succession, the probability that the presentation of the first on a later occasion will lead to recall of the second is assumed to be directly related to the degree of similarity of context on the two occasions.

The third factor has to do with the fact that if the stimulus S_1 has occurred in contiguity with the stimulus S_2 on one occasion and with a stimulus similar to S_2, say S_2', on another occasion, then upon later presentation of S_1 there is uncertainty as to whether S_2 or S_2' will be recalled. This type of interference in memory can be overcome only by constructive intellectual activity on the part of the learner which leads to selective storage of associations which avoid the interference.

For example, suppose the stimulus combination AB has been followed by C on one occasion and the combination AD has been followed by E on another. Confusion in recall will be avoided if the learner is led to attend selectively to the unique components, the B and D, and to store in long-term memory the associations BC and DE, thus avoiding errors that might arise from associations of component A with either C or E. The learner can achieve this result if, for example, upon presentation of the combination AD-E he recognizes the ambiguous associations of component A with C and E and thus rehearses only the association D-E following the second presentation.

The basic condition for transfer of learning from one situation to another is assumed to be communality in constituent associations. If events A and B have occurred in succession and have become associated in memory, then a representation of B will tend to be reinstated in memory on later occasions by other events which are similar to or overlap with A in proportion to the degree of similarity or overlap. It is assumed that once associations have been formed and consolidated this transfer occurs in a largely automatic manner. However, various activities on the part of the learner during a learning experience can influence transfer to an important extent. On the one hand, as illustrated in the preceding paragraph, undesired transfer can be reduced if the learner is led to attend selectively to stimuli which are unique to a given situation and to rehearse only associations involving these unique aspects. On the other hand, transfer can be enhanced if the learner attends selectively to stimuli or stimulus properties which are common to the original and the transfer situations and selectively rehearses these.

7. probability learning

Although within a more formal statement of the theory sketched here, probability learning can be derived from other, more elementary mechanisms, we shall simply state it as an independent assumption that the organism stores information with respect to relative frequencies of sequences of events. If, for example, event A has been followed by B and C on different occasions in the past, then the organism's tendency to anticipate B following the occurrence of A is directly related to the relative frequency with which B has actually followed A.

Probability learning provides the basis for the continual modulation of behavior by variations in anticipation of reward and punishment. When exposed to an environment containing stimuli which have had different past histories of association with punishment and reward, the organism receives positive feedback on the average whenever it shifts attention from a stimulus which has had a lower, to one which has had a higher relative frequency of association with reward and, conversely, receives inhibitory feedback whenever it shifts attention from a stimulus more highly correlated with reward to one less highly correlated with reward in the past.

8. formation of hierarchies of stimulus and response units

A major consequence of the associative process is that the learning organism continually builds higher-order units, generating the phenomena termed stimulus patterning and response integration. If a number of originally unrelated stimuli repeatedly occur together, multiple associations form among the organism's reactions to the various components, so that ultimately the occurrence of any component reinstates a representation of the whole set in memory; once this learning has occurred, the set of stimuli enters into association with other behavioral events as a unit. One describes the end result of this learning by saying that the organism has come to respond to the collection of stimuli as a pattern.

Similarly, members of a set of motor reactions which are repeatedly evoked in close contiguity come to be associated so that evocation of a component reinstates a representation of the entire set, which then may enter into association with other behavioral events as a unit. The former aspect of unit formation is most conspicuous in learning to read, when letters, syllables, and words come to act in certain respects as unitary patterns. The latter aspect is similarly conspicuous in learning to speak and in the development of verbal control of behavior as, not only words, but propositions expressing rules and strategies come to act as units in the continuing elaboration of the associative learning process.

Organization of still higher-order stimulus–response units occurs in discriminative and conceptual learning, and in free recall. Discrimination problems are frequently generated by combining stimulus dimensions, as defined by the experimenter, in various ways and arranging contingencies of rewarding and nonrewarding outcomes upon response to particular dimensions or combinations of dimensions. Once the organism has learned to associate a common perceptual or mediating response with the various values on a stimulus dimension, for example colors or forms, it is in a position to learn to approach or avoid entire dimensions in new problems, thus greatly short-cutting the process of learning to approach or avoid the various constituent stimuli. Similarly, sets of responses, such as names of colors, names of states, or the like, can be associated with common mediating, or categorizing responses; then in new situations the individual's capacity for recall can be greatly amplified. Suppose, for example, that the list of words read to a subject in a free-recall experiment included the first ten digits, the primary colors, and the names of the New England states. A sufficiently sophisticated learner would need only to rehearse and transfer to long-term memory the category names, or code designations, of the three classes in order to be prepared to reproduce the entire list on request.

9. learning to learn

Beyond the broader developmental changes in learning abilities treated in the following section, more localized changes in learning rates are generally observed as a function of experience with particular types of tasks or materials. Within the theory outlined here, these latter, more specific, types of improvement in learning rate are derivable from other aspects of the theory and do not require independent assumptions. Nonetheless, because of the special importance of "learning to learn," or the development of learning sets, for interpretation of mental development and retardation, it will be useful to summarize the principal concepts involved.

Many cases of apparent learning to learn represent, not changes in learning rates, but simply transfer from learning which has occurred in situations involving common associative relationships. An elementary example is the progressive increase in rate of acquisition during a sequence of acquisition and extinction sessions in classical or operant conditioning (for a review of the literature, see Capaldi, 1967). As a result of learning in earlier sessions, the unconditioned or reinforcing stimulus itself becomes a signal for further reinforcement; since this component of the stimulus complex which normally precedes reinforcement during acquisition is necessarily absent on extinction trials, and thus is protected from extinction, its effectiveness is immediately reinstated upon the first trial of reconditioning. This type of increase in learning rate by virtue of specific transfer is especially pervasive in the learning of language. Associations formed to the roots of words, for example, and to various prefixes and suffixes contribute directly to appropriate performance upon the occurrence of new words involving common elements.

Whereas the former examples involve specific transfer of associative relationships between the response being learned and elements of the transfer situation, progressive increases in learning rate in discrimination and conceptual learning situations are probably more often mediated by observing or attentional responses to cues or stimulus dimensions which are common to broad classes of problems. Over a series of form-color discrimination problems, for example, the individual learns that attention to the relevant dimension is associated with reward whereas attention to irrelevant dimensions, such as size, is not, so that in later problems of the series attention is confined to relevant stimulus dimensions from the beginning and "unnecessary" errors are eliminated. Similarly, in learning to read the child learns to attend to the forms of letters rather than to size of print, or the like.

Even when successive discrimination problems are unrelated with regard to stimulus relationships, the initially naive organism will exhibit a progressive increase in acquisition rate simply as a consequence of learning that premature choice responses are associated with higher probabilities of nonreward than delayed choice responses, thus coming to inhibit overt response on any trial until all of the stimuli in the display have been scanned.

In verbal learning and recall, in addition to all of these more elementary mechanisms, improvements in learning or retention of various degrees of specificity may result from the acquisition of rehearsal and organizing strategies. With regard to the former, over a series of similar tasks the individual may learn which particular stimulus properties or relationships need to be selectively rehearsed in order to lead to increased probabilities of reward. With regard to the latter, various types of categorizing and grouping strategies, usually discovered by trial and error, may be rewarded and then facilitate recall performance over a series of related tasks.

It will be observed that, according to the present analysis, the term "learning to learn" is in a sense a misnomer. An individual who has had considerable experience with a particular type of task does not actually learn more rapidly in the sense of accelerating the formation of new associations; rather, he needs to learn much less in order to attain a given level of performance.

II. *Interpretations of Standard Learning Situations*

In this section we review several standard learning experiments, identifying basic processes within the framework of the theory sketched above and attempting to identify important sources of individual differences in learning ability.

1. classical conditioning

Pavlovian conditioning is perhaps the simplest of the standard learning experiments in that the course of conditioning may be relatively free of specific dependence upon previous learning. Nonetheless, even in this apparently highly simplified experimental paradigm, there is room for multiple processes. And complexities multiply rapidly with even seemingly innocuous modifications of classical experimental arrangements.

Let us consider, as a specific example, conditioning of the eye blink with an electric shock to the cheek as the unconditioned stimulus (US). Conceptually, this process is somewhat simpler than the more frequently studied variant with an air puff to the cornea as the US or than salivary conditioning with acid as the US. The simplification arises from the fact that the occurrence of the conditioned response (CR) has, so far as is known, no appreciable effect upon the individual's reaction to the US. Prior to conditioning, the shock elicits a defense reaction, including the eye blink and the activation of a negative drive amplifier which provides facilitative input to the family of defense reactions. The conditioned stimulus (CS), usually a flash of light or the sound of a buzzer, normally evokes an original reaction of which an eyelid response is only a minor component. Over a series of presentations of the CS followed by the US, associations form between representations of these stimuli, but owing to random fluctuations both in the environment and in the state of the subject only a sample of the population of available associations is formed and consolidated on each trial.

The first association formed is sufficient to lead, upon occurrence of the CS, to anticipation of gross aspects of the unconditioned reaction (UR), including in particular the negative drive input. Frequently after only one conditioning trial, the individual will yield evidence of anticipating the US upon presentation of the CS. However, a considerable number of trials may be required before enough associations are formed and consolidated so that a CS presentation brings into immediate memory a sufficiently detailed representation of the US to lead to an anticipatory occurrence of the particular response pattern, in this instance the eye blink, which meets the criterion for a CR.

A principal reason why individuals of very different learning ability generally exhibit little if any difference in rate of simple CR acquisition may be the fact that in standard experimental procedures many sources of individual differences in conditioning rate are minimized. Individual differences in intensity or extent of response to the US are generally compensated experimentally by adjusting the intensity of the US for each subject so as to yield URs of approximately the same magnitude. Similarly, individual differences in the tendency to observe or attend to the CS are minimized in standard experimental procedures by measures which are taken to insure adequate stimulation of the subject by the CS on each trial. One might suppose that there would be major differences in the number of associations formed per trial when other factors are held constant; however neither this supposition nor the related expectation that learning ability should be intimately tied up with rate of formation of associations has found substantial experimental support. Measures which reduce individual differences in reactions to the CS and US have been found to reduce individual differences in observed rates of acquisition, sometimes almost to the vanishing point (see, e.g., Voeks, 1954).

Why should individual differences in rate of extinction of a CR be more systematically related to other measures of learning ability than individual differences in acquisition? The answer may lie in the fact that the extinction procedure is less well controlled. In part, extinction involves a new conditioning process in which the individual learns to anticipate absence of the US following CS presentation. This event, however, is less well defined than US occurrence, and there is no way (comparable to adjusting intensity of a US) by which the experimenter can compensate for individual differences in responsiveness to the changed state of affairs following the CS. At the time of occurrence of a CS on a late acquisition trial, the individual may be influenced by the stimulus traces from the URs of the previous trials or by representations of these in immediate memory. However, immediately following the first extinction trial this source of contextual stimulation is no longer available. The change in context could lead to a decrease in the individual's tendency to anticipate the usual consequence (that is, the occurrence of the US) even prior to any new learning (counterconditioning).

The importance of these changes in context depends to a major extent on the individual's attentiveness to background changes. To the extent that the influences of changes in context are mediated by short-term memory processes, and especially by verbal processes, they would tend to be related to mental ability measured by tests involving these processes.

Some implications of the present interpretation are readily testable. For example, although rate of initial acquisition of CRs has been found to vary little over a wide range of IQ levels, we predict that if the same individuals were given a series of repeated acquisitions and extinctions of the same CR in a given situation, the progressive increases in acquisition rate routinely observed under these conditions would be greater for individuals at higher IQ levels. Therefore a relationship between rate of conditioning and mental ability would appear with experience even under conditions such that none would be manifest during initial acquisition.

This prediction is based on the idea that the same types of change in background, or stimulus context, which contribute to rapid extinction following initial acquisition would operate in reverse, so to speak, to facilitate reacquisition. The individual who is sensitive to the more remote stimulus relationships over sequences of trials may learn that the occurrence of a US on any trial of a series greatly increases the probability that it will occur on the next; thus, with experience over a sequence of acquisition and extinction blocks, such an individual might come to exhibit relatively high rates of both conditioning and extinction.

Whereas the number of temporary associations established on a conditioning trial evidently involves sufficient multiplicity and redundancy so that individual differences in this respect are not a major source of variance in conditioning rates, one might suspect that an important locus of individual differences would lie in the "triggering" of the process of consolidation or transfer to long-term memory storage, or in the susceptibility of this process to disturbance by events occurring during a short interval following a conditioning trial. However, although conditions influencing the course of consolidation have begun to receive rather intensive study (see, e.g., Grossman, 1967, pp.

797-814), there seems to be no information available concerning individual differences in this process or relationships to standard measures of mental ability.

One must emphasize especially that, since the eye blink is under voluntary control in normal individuals, and doubtless in all but the most severely retarded, it is only under especially simplified conditions that one can expect to study conditioning uncomplicated by more complex mental processes. Adequate simplification is probably approximated only when a US, such as shock to the cheek, is used which is not modified appreciably by occurrence of the CR and when during the experiment the subject is engaged in some task which occupies his attention and minimizes the likelihood that verbal expectations concerning occurrence of the US will modify his response to the CS.

When a US such as an air puff to the cornea is used, the situation necessarily becomes more complex, since occurrence of the eye blink during an appropriate interval following onset of the CS reduces the unpleasant stimulation from the air puff, thus establishing avoidance contingencies. In this case the individual is in a position to learn a simple discrimination, namely that occurrence of the CS plus the eye blink leads to less unpleasant stimulation of the cornea than occurrence of the CS and omission of the eye blink.

Acquisition of this discrimination is of course facilitated by verbal rehearsal of the relevant stimulus relationships, and thus it is not surprising that, in the standard eyelid conditioning situation with air puff as the CS, rate of conditioning is considerably higher for college-student subjects than for retarded groups (Ross, 1966). It remains to be determined whether the conditioning rates for the retarded could be brought up to the normal level in this situation by appropriate instruction and training with respect to attention to the stimulus contingencies and rehearsal of stimulus relationships.

2. operant conditioning

Superficially, operant conditioning might seem simpler than classical conditioning, since it does not involve the transfer of a response from one stimulus to another but only the strengthening of a response which the individual already has some tendency to make in the given situation. Within the present theoretical framework, however, it is recognized that response tendencies are not simply strengthened in a direct way by rewarding consequences. Rather, a rewarded occurrence provides an opportunity for the individual to learn associations between stimuli encountered in the course of executing the given response and the rewarding consequence.

In an operant situation the learner is not constrained to attend to particular stimuli at particular times as in the classical conditioning experiment, and at the beginning of training he first receives reward when he happens to run through the prescribed stimulus-response sequence in the course of more-or-less random activity. As a result of this experience, learned associations may be formed between the successive behavioral events in the sequence, the likelihood of this occurring being increased to the extent that the individual attends to the relationships between successive stimuli and rehearses these following the receipt of reward. Once the associations have formed, the individual upon later encountering the first stimulus in the sequence, will be led to anticipate reward, thus generating facilitative feedback and increasing the likelihood that the stimulus-response sequence leading to reward will run to completion.

Although there are doubtless individual differences in rates of formation of associations and transfer from short- to long-term memory storage, it seems likely that, to even a greater extent than in the classical conditioning situation, individual differences in rate of operant conditioning are associated largely with differences in tendencies to attend to relevant stimuli or aspects of the stimulating situation.

In the present theoretical framework it is assumed that the mechanism producing facilitative feedback is activated by increases and inhibited by decreases in anticipation of reward. In order for an individual's behavior to be brought under the control of feedback from anticipation of reward he must learn that changing a stimulus situation in one direction (for example, by approaching a lever) increases the probability of reward and changing it in the other direction (by moving away from the lever) reduces probability of reward. This kind of control is not obtained if, for example, the individual makes the response which produces reward while attending to an irrelevant aspect of the stimulus situation.

In the training methods used by students of operant conditioning, much more care is generally taken in arranging contingencies between responses and rewards than in insuring that the learner will have an opportunity to observe relationships between the stimulus changes and probability or immediacy of reward.

The importance of stimulus variables is seen especially clearly in the

S.C.E.R.T., West Bengal
Date 15-1-72
2999
154·4
EST
S.C.E.R.T., West Bengal
Library

studies of operant conditioning of verbal behavior (reviewed by Krasner, 1958). In free-responding situations, the subject produces a continuous stream of speech, and utterances belonging to some particular class receive reward. Opportunities for the subject to perceive relationships between stimulus changes and reward probabilities are slight, and characteristically, rate of conditioning is very low.

By contrast, when the situation is modified so that each successive trial begins with the display of a set of stimuli, response to one of which eventuates in reward, conditions are relatively favorable for the perception of stimulus-reward relationships, and rate of conditioning is relatively high (Taffel, 1955). Thus over the wide range of operant conditioning situations in which stimulus variables are relatively uncontrolled, we should expect wide individual differences in observed rates of conditioning to result from individual differences in tendencies to utilize observing or attentional responses with respect to stimulus relationships.

Even more than in the case of the simple operant responses usually studied in the conditioning laboratory, acquisition of the finely differentiated response patterns involved in perceptual-motor skills depends to a major extent upon the learning of relationships between stimulus changes and changes in probability of reward. To take a simple example, if a child is being trained to touch with a pointer a spot of light which may appear anywhere upon a screen, learning, if manifest at all, will be painfully slow if a separate stimulus-response association has to be formed for each possible location of the spot. Acquisition of the skill depends upon learning that reduction in the distance between the pointer tip and the light spot, regardless of the location of the latter, signifies an increase in probability (or reduction in delay) of reward and that, contrariwise, increase in this distance signifies a decrease in probability of reward. When this relationship has been learned, the child's movement of the pointer is continuously modulated by the visual feedback from the changing distance between the pointer and the spot.

We must expect major individual differences in speed of acquisition of perceptual-motor skills to be related to differences in attentiveness to relevant stimulus dimensions and, similarly, major differences in efficacy of different types of training procedures to be related to the extent to which they successfully direct the learner's attention to the relevant stimulus relationships.

3. discrimination learning

In modern learning theories discrimination learning is generally assumed to involve the same basic processes as simple classical or instrumental conditioning. Particular theories differ largely with respect to assumptions as to how more elementary processes are combined or organized in discriminative situations. Within the theoretical framework developed in the previous chapter, a discrimination problem can be analyzed in terms of a set of possible stimulus-response sequences, each beginning with attention to some aspect of the stimulus situation present on the trial and eventuating in a rewarding or punishing outcome.

The learning which goes on in the course of mastering a discrimination problem is a matter of learning the relationships between the successive behavioral events in these stimulus-response sequences. The process involves the establishment of temporary associations, simply as a result of contiguous occurrences of events, followed by the consolidation, or transfer to long-term memory, of subsets of these associations as in simpler conditioning situations. Owing to the structure of the discrimination problem, stimulus elements which initiate some sequences will, in the course of learning, come to serve as signals for anticipation of reward whereas others will become signals for nonreward.

On the average, discriminative performance is expected to improve as learning proceeds, but the relationship is by no means one to one. At the beginning of learning, performance with respect to the stimuli to be discriminated is generally random. At the end of acquisition, attention to various aspects of the stimulus situation leads to different degrees of anticipation of reward and consequently different degrees of facilitative feedback to behavior sequences which have followed these in the past. As a consequence, stimulus-response sequences having the greatest likelihoods of leading to reward tend to be selected in the presence of each of the stimulus situations involved in the problem.

Improvement in correct responding may be expected to occur primarily at points in the learning sequence where the individual shifts attention from cues which are unrelated to cues which are more closely related to reward and nonreward, that is, from irrelevant to relevant components or properties. During sequences of trials when the learner is attending only to irrelevant cues and basing his responses upon these, there will be no observable change in probability of correct responding. Therefore, as emphasized by Zeaman and House (1963) in their analysis of discrimination learning, major sources of individual differences in rate of discrimination learning must be differences in initial tendencies to sample, or attend to, different aspects of the stimulus situations involved, and also individual differences in sensitivity to the different degrees of anticipation of reward.

Doubtless, individual differences in both of these respects are influenced to an important extent by learning which occurs in many different situations in the course of normal development or in the course of specific training in a particular type of situation. The former aspect has received considerably more attention in the literature (see, especially, Zeaman & House, 1963) than the latter. Nonetheless, it seems clear in terms of the present analysis that an individual's level of anticipation of reward upon entering a new discrimination problem should be an important factor in rate of learning. If, for example, an individual is accustomed to failure in any given type of discriminative situation, then upon facing a new problem his a priori anticipation will be only 50% reward. Stimulus-response sequences which eventuated in 50% reward (that is, those arising from attention to irrelevant aspects of the situation) yield the rate of reward he has come to expect and consequently there is no reason for him to deviate from his habitual mode of response. However if the individual is accustomed to success, then upon confronting a new problem the 50% frequency of reward associated with response to irrelevant features of the situation represents a distinct reduction from his a priori expectations; thus ongoing behaviors will tend to be inhibited by reduced facilitative feedback, permitting him to sample other stimulus-response sequences which may lead to reward with higher probability.

The importance in discrimination learning of various general habits which carry over from one problem to another is so great that the differences in rate of discrimination usually observed between experimentally naive individuals of different ages on IQ levels, or even members of different species, can often be reversed by appropriate training. Thus even the rat learns with experience in a two-choice discrimination situation to slow down at the choice point and examine the stimulus display fully before making its overt choice, and having done so, comes over a series of reversals of a discrimination problem to learn each new discrimination in a single trial (Estes & Lauer, 1957).

Although direct evidence is scanty, there are both theoretical and empirical grounds for the supposition that short-term memory enters into discrimination learning in a more complex way than into simple conditioning. Although, on the whole, discrimination learning, like most other forms of learning, tends to be facilitated by distributed practice, a number of studies provide rather clear indications that the relationship is reversed at long intervals, with an optimum intertrial interval in the neighborhood of 3 or 4 min for visual discrimination learning in the rat (Pennington & Thompson, 1958; Weinstock, 1958).

A theoretical interpretation of the optimum interval can be offered in terms of two processes that would be assumed to be operative within the present theoretical framework. First, if successive trials are too close together, the temporary associations established on one trial will have little or no opportunity for consolidation, or transfer to long-term memory, before onset of the next trial sets up interfering associations; this process in itself would yield a simple, monotone function, relating speed of discrimination learning directly to intertrial interval.

In another respect, however, relatively short intertrial intervals would be advantageous for discrimination learning. If the intertrial interval is such that the event sequence of one trial is still fully represented in short-term memory when the next trial occurs, the learner will be in a position to notice relationships between changes in the stimulus and changes in the rewarding outcome which otherwise would be likely to escape attention. Thus, for example, Weinstock (1958) found that rats showed no progress whatever toward learning a discrimination between a flickering and a constant light over a large number of trials when the intertrial interval was 24 hr but that they learned the same discrimination rapidly when the interval was reduced to 5 min.

My interpretation is that the animal has very low probability of attending to either condition of illumination unless trials are close enough together so that it notices the change from one to the other. Presumably, other things equal, the probability of attending to changes in the stimulus situation would be a monotone decreasing function of intertrial interval. The combination of the two gradients would yield a nonmonotone function for rate of discrimination as a function of intertrial interval, with an optimum at some intermediate value, in apparent accord with the combined results of a variety of studies in this area.

The importance of the second of these two processes may be much greater in human beings than in animals, since by means of verbal rehearsal the human learner can greatly extend the effective short-term memory function. We should expect both that individual differences in human discrimination learning would be related to rehearsal tendencies and also that deficiencies in discrimination learning could be substantially modified by special training which leads to the establishment of effective rehearsal strategies.

4. learning set

The term "learning set" was introduced by Harlow (1949) to refer to a specific form of learning to learn observed in visual-discrimination studies with monkeys. In the standard experiment the animals were given series of similar but formally unrelated problems (unrelated in the sense that solution of any one problem gave no information concerning the solution of others in the series). Progressive improvement was observed in the rate of learning these discriminations; this rate of improvement defines the concept of learning set. Analogous experiments were conducted later in paired-associate learning situations with human subjects by Greenberg and Underwood (1950) and by Meyer and Miles (1953).

What is the theoretical basis for the progressive improvement in discrimination learning ability exhibited in these learning set studies? Any type of theory can accommodate the phenomenon of progressive interproblem improvement, but only certain types of theories provide a positive basis for expecting it. The classical association or conditioning theories of, for example, Hull and Spence, or basic stimulus sampling theory can interpret the phenomenon by assuming changes in parameter values.

But what processes should be expected to lead to progressive changes in parameter values as a function of experience? One possibility lies in such general strategies as slowing down at the choice point before selecting a response (Estes & Lauer, 1957). Within the type of discrimination learning theory developed by Zeaman and House (1963), there is a more constructive basis for interproblem improvement. The same process of learning attentional responses to relevant dimensions which is assumed to be involved in the learning of any one discrimination and which accounts for the advantages of intradimensional shifts would imply progressive improvement in discrimination learning rates over a series of experiences with problems involving the same dimensions.

House, Smith, and Zeaman (1964) have pointed up the paradox which is to be found in the results of studies of learning set, on the one hand, and of proactive inhibition, on the other. According to the former, an individual comes to learn more and more rapidly as a function of experience with a particular type of material, but according to the latter, he comes to forget more and more rapidly what he has learned as a function of the same kind of experience. Both processes were brought out nicely in the study of paired-associate learning reported by House, Smith, and Zeaman. The subjects were mentally retarded children with mental ages (MAs) of 4-8 years who learned a succession of paired-associate lists with retention tests averaging about 40 hr after learning. The geometric mean number of errors to criterion decreased from 4.0 on the first list to 3.5 on the fourth, 3.0 on the seventh, and 2.1 on the tenth. Concurrently, percentage recall decreased from 86 following the first list to 56 following the fourth, 53 following the seventh, and 43 following the tenth.

An analysis of errors on both learning and recall tests was instructive, showing that during original learning of each list, a very large proportion of errors were intralist intrusions, whereas on recall tests a very large proportion of errors were interlist intrusions. Thus, during learning errors were primarily the result of failures to differentiate among the items within a list and the learning-to-learn effect represented a facilitation of this process. Conversely, errors of recall were primarily a matter of confusing items between different lists, and the frequency of such errors increased with the number of lists that had been learned. It should be noted that there is a basic incompatibility between the most efficient method of coding the items in a particular list in order to facilitate learning and the most effective method for maintaining recall of a large collection of lists.

These results and related ones suggest that the disappointingly poor long-term retention frequently found for school learning may be a result of teaching methods which emphasize efficient learning of small blocks of material within particular class sessions at the cost of measures which are needed to maintain long-term recall. Some of the measures needing investigation with a view to improving long-term recall are modification of coding strategies to reduce interlist interference, techniques for promoting organization of recall, and the use of periodic tests and rehearsals.

5. paired-associate learning

In terms of the experimental arrangements, a paired-associate task simply involves the learning of a number of concurrent discriminations. In practice, the materials used in paired-associate learning experiments are largely verbal and this factor greatly increases the scope for transfer and learning-to-learn effects. If a single paired-associate item, say the relation between an English word and the corresponding word in a foreign language, were considered entirely in isolation, the processes involved in acquisition would presumably be essentially the same as those in classical conditioning. On any one training trial, representations of the stimulus and response members of the item and associations between them would be established in the temporary memory system and over a subsequent interval of time, in the absence of interference, transfer of a portion of this information to long-term memory would occur. With new samples of associations being established on successive trials, the result of a sufficiently long series would be that the presentation of one member of a pair would evoke in memory a relatively full representation of the other.

However, it is likely that, in adult human beings, paired associates are rarely learned in this fashion. Frequently the process can be greatly short-cut in that the learner may find some feature or component of the stimulus member of a pair which, through some chain of existing associations, brings to mind features of the response member. Indeed, in many cases the chain of associations is found to include the name of the response member, in which case learning may be complete for practical purposes in a single trial.

Thus individual differences in the observed rate of learning of single paired-associate items must be expected to depend to a major extent upon the amount of previous learning of similar material and upon previously established habits of looking for mediators. However, neither in laboratory experiments nor in most practical situations, such as the acquisition of languages, are single paired-associate items learned in isolation. Generally the task involves the learning of a list, in which case the amount of information needed to provide a basis for uniform correct response depends upon the number of items in the list and the amount of communality between items.

In general only some subset of the components or aspects of each stimulus member of the list need be represented in memory in order to differentiate all of the stimuli from each other. Consequently, for any given list, there is some minimal set of stimulus elements or aspects which must be represented in memory in order to permit 100% recall. The number of associations which must be established during the learning of the list thus depends to an important extent upon the learner's ability to discern or attend to features which differentiate among the items of the set. By selectively attending to and rehearsing relationships involving these distinguishing features, an individual can greatly reduce the amount of learning that needs to be done before he can meet any prescribed criterion of correct performance. The bright college student will typically master in a very few trials a paired-associate list which might be completely beyond the capacities of a retarded child, not because the college student's rate of formation of associations is very different but primarily because he needs learn so much less in order to achieve correct performance on the same set of material. When the paired-associate situation is especially simplified so as to reduce the possible contribution of these various auxiliary processes, differences in acquisition rate between individuals of different IQ levels should diminish greatly. There seems little room for doubt that the inefficient performance of the mentally retarded individual in verbal paired-associate learning reflects to a major extent his retardation in the development of various habits of selective attention, search, and rehearsal. Little seems to be known, however, concerning the extent to which his retarded development is in turn a function of impoverished opportunities to learn relevant habits in earlier life or the extent to which the retardation might be overcome by special training.

III. *Individual Differences and Developmental Processes*

1. the concept of learning ability

How are learning abilities to be conceptualized within the frame of reference provided by modern learning theory? First of all, it is clear that the notion of learning ability must be abstracted from performance in particular learning situations. Substantial individual differences with respect to any measure of speed of learning are universally observed when a group of individuals, however similar on any selection criteria the experimenter may use, are run through an identical experimental routine. And this is true whether the situation be as simple as classical conditioning or as complex as classroom learning. If the individuals are well matched with regard to whatever previous experience is relevant to the particular task, then these differences in observed rates of learning are attributed to differences in learning ability.

There is no reason, however, to assume anything like a one-to-one relationship between abilities and types of learning tasks. Rather, there are substantial grounds for assuming that differences in observed learning rates in any given situation are a joint resultant of individual differences in elementary capacities and in cumulative results of past learning in the same and other situations. The balance of contributions from these sources must be expected to vary from one type of learning situation to another.

By "capacity" I refer to any source of individual differences, determined either genetically or by some combination of genetic factors and developmental processes which are independent of previous learning. It is difficult to be more specific, for to date little progress has been made toward identifying particular learning capacities. Their existence is inferred indirectly from patterns of differences in observed performance and from analogies to better-understood classes of behavioral phenomena, for example, those having to do with sensory and motor processes.

Although it is not yet possible to point to any relatively elementary learning capacities, comparable say to those having to do with visual or auditory acuity in the sensory sphere, there are indications that some identifications may be realizable within the foreseeable future. For example, in the study of phenomena of memory among individuals with neurological disorders, some relatively clear cases have been reported of individuals who retain normal performance on immediate-memory tests but exhibit complete loss of the ability to consolidate material which has been entered in immediate memory (Milner, 1966; Wickelgren, 1968). But, if transfer of information from short-term to long-term memory depends upon a relatively specific neurophysiological mechanism which can be selectively eliminated by neurological damage, then it seems most likely that there will be found to be individual differences in the efficiency of operation of this mechanism and that these will in turn be found to make some contribution to efficiency of performance in tests of memory and in learning situations.

Primarily on the basis of analogy to better-understood systems, it seems likely that the maturation of the anatomical and physiological substrate of elementary learning capacities must be essentially complete at a very early age in human beings, quite possibly within the first year of life, and that individual differences in rates of maturation can have little to do with individual differences in the development of learning ability from infancy to adulthood. Differences in capacities may well be involved in mental retardation; but if so, it would seem that the critical differences must lie primarily in limits rather than rates of development. Thus "mental deficiency" may be a more appropriate term than "mental retardation."

There are abundant reasons to believe that even if the isolation of elementary learning capacities and their genetic determiners can be achieved, we will come into possession of no magic key to the understanding of the normal development of learning abilities or the modification of deficiencies. Except perhaps in infancy, a great part of the variance in rates of learning between individuals must be attributable to differences in the results of past learning. Observed rates of learning in most situations, we now realize, depend to a major extent upon habits or strategies of selective attending, seeking of information, coding and recoding of stimulus information, rehearsing, and the like, and the manner in which these are organized. Individual differences in these habit systems may indirectly reflect differences in capacities, but they must also be strongly determined by variation in motivational systems and in previous opportunities to learn. Thus, Gagné (1968) speaks of the "cumulative-learning" hypothesis of intellectual development. New learning at any age depends to a major extent on the recombining of previously acquired discriminations, behavior sequences, principles, and concepts.

2. development of learning abilities

The differences between the infant and the adult human being with respect to the results of experience, that is, the accumulation of knowledge and skills, are, at the same time, so vast as to defy description and so obvious as to be taken for granted. But only the deepened understanding provided by decades of research has begun to make it clear that differences in learning abilities between infant and adult are of at least the same order of magnitude and are at present far beyond our ability adequately to conceptualize. Nonetheless, once our observation of learning moves from the uncontrolled situations of everyday life into the laboratory, we cannot help but be almost equally impressed by the similarities between infant and adult learning.

Under experimental conditions which minimize the role of voluntarily controlled behavior, the acquisition of conditioned responses proceeds at very similar rates in animals, human children, and human adults. Similarly the rate of paired-associate acquisition for simple, nonverbal stimuli and responses varies little with age over most of the life span and differs little between mentally retarded and normal groups. And, retention of simple associations once established, if not actually independent of age and other indices of mental development, at least exhibits no systematic variation that has been detectable by research to date.

To my knowledge the only serious effort toward a theoretical interpretation of this paradox is that presented by Hebb (1949). I shall assume a familiarity on the part of the reader with Hebb's analysis, though in the present review there will be no occasion to go into Hebb's neurophysiological model. To almost the same extent as was the case when Hebb was developing his pioneering formulation some 20 years ago, an approach to developmental phenomena in terms of learning theory today must be largely speculative. The enormous literature on changes in sensory, motor, and intellectual performance with age is still accompanied by only the scantiest beginnings of a parallel accretion of factual information concerning developmental changes in detailed aspects of learning. Useful descriptions of the stages of intellectual development characteristic of children within certain cultures have been provided, notably by Piaget (Piaget, 1926; Woodward, 1962), but there remains the task of interpreting the sequences of stages in terms of learning processes.

Associative Processes. With regard to the most elementary aspects of the associative process, the preponderance of evidence indicates that the rate of establishment of temporary associations between contiguous occurrences of behavioral events varies extremely little over the lifetime of the individual and, further, is extremely resistant even to brain damage. In the senescent normal individual and in the mentally retarded (see, e.g., Ellis, 1968), very-short-term memory of the type presumed to be mediated largely by these temporary associations appears relatively unaffected.

Quite the contrary is the case with regard to consolidation of these associations, or the transfer of the information contained into long-term memory, a process which occurs at rates that change with age in a curvilinear fashion. The rate of this transition process presumably increases during early development, though explicit documentation is difficult to find, and certainly decreases with the onset of senility. Also it is extremely sensitive to disturbances of the nervous system and to many types of brain damage. A notable example of the dissociation of the processes of establishment and consolidation of temporary associations is observed in the neurological disorder termed Korsakoff's syndrome.

One source of this disparity may lie in the vast multiplicity of elementary associations set up by any one learning experience. The degree of redundancy is probably so great that substantial variation in numbers of temporary associations established might have no detectable effect on short-term memory. On the other hand, there is considerable reason to believe that the process of consolidation, or transition to long-term storage, though doubtless modifiable by many conditions, may be initiated following any given learning experience in an all-or-none fashion by some mechanism, as yet unidentified, which is relatively localized. Lines of relevant evidence range from the indication that consolidation of conditioned responses in gold fish is initiated only upon removal from the apparatus (Davis, Bright, & Agranoff, 1965) to data on so-called stimulus selection in paired-associate learning (Trabasso & Bower, 1968), and to demonstrations of "blocking" of conditioning by redundant stimuli (Kamin, 1967).

The changes in the neurophysiological mechanism responsible for initiating consolidation which occur in the course of maturation and aging may well be largely independent of external control. At the same time, there is no doubt that operation of this mechanism may also be influenced to an important extent by habits of rehearsal and tactics for protecting newly formed associations from interference. Most attention tends to be given in the literature to the long-term possibility that functioning of consolidation or storage mechanisms might be improved by drugs or other physiological manipulations. But the more immediate fruits of research in this area, which might yield tangible improvements in learning efficiency on the part of normal individuals as well as the retarded, would seem to be the habits of selective attention and rehearsal and other behavioral tactics that can be developed by training and brought under voluntary control.

The other side of the short-term retention process perhaps merits more attention than it has had in the literature. Efficient performance in many situations may depend as much upon efficient short-term forgetting as upon long-term retention. Consider, for example, the plight of the operator of a calculating machine if he were unable to clear from memory the input from each successive display on the keyboard. In this type of situation, retention may be prevented to a considerable extent by the individual's switching attention rapidly from one display to the next, thus minimizing the interval free of interference which is evidently needed for the initiation of the consolidation process with respect to any one input.

Although there seem to have been no systematic empirical studies of developmental changes in the efficiency of "short-term forgetting," indirect evidence from work of Spitz and his associates (1959, 1969) suggests that there may be significant differences between retarded and normal individuals in the speed with which shifts of attention occur. Unfortunately, at present there seems to be no basis whatever for judging to what extent the observed differences reflect retarded development of some as yet unidentified neurophysiological processes or indirect effects of other developmental processes which may depend, at least in part, on previous experience.

Formation of Higher-Order Behavioral Units. The major source of the difference in rates of conditioning or associative learning generally observed between immature and mature organisms is simply that the former have more to learn in order to meet the criteria for learning in any given situation. Very early in development, if each of two events which occur in contiguity comprises a number of aspects, associations among all of the components must be formed and consolidated before the learning process can be considered complete. Suppose, for example, an event with components A, B, and C were followed by an event with components D, E, and F. Temporary associations would be set up between components of the first event and components of the second. However, in the absence of rehearsal, only some random sample of the associations formed on the given learning experience would subsequently be consolidated. Since the elementary associations fluctuate between active and inactive states over time, upon the next occurrence of the first event some of the associations formed on the first occasion might be inactive. Owing to these two factors, repeated trials would be required before enough associations could be formed to provide a basis for reliable recall of the second event upon occurrence of the first. If, however, in the course of the organism's development these events were to recur many times, whether together or separately, associations would form among the components of each, so that ultimately the occurrence of any component of either event would tend to reinstate in memory a representation of the whole of that event. If the two events first occurred in contiguity after each had taken on such a unitary character, one might observe "one-trial learning." Associations between a constituent of the first event and any of the constituents of the second would be sufficient, upon the recurrence of the former, to reinstate in memory a representation of the latter.

Hebb (1949) first assembled the empirical evidence and offered a neurophysiological interpretation of the important developmental period of "slow learning." During this period recurring perceptual and motor patterns become integrated through multiple associative learning into higher units, thus preparing the way for the facile, all-or-none learning characteristic of the mature individual. Analyses of factors involved in the development of "reading readiness" have demonstrated progressive improvement over the ages of four to eight years in ability to recognize transforms of critical features of stimuli, for example, letter-like forms, which appear in printed material (Gibson, 1965). Impoverishment of a child's early environment, including especially any major restrictions on play activity or lack of feedback from older individuals with respect to vocalizations, must be expected to retard severely this important process of development of behavioral units and thus in turn to restrict the range of situations in which the individual could be expected to manifest normal learning rates in later life.

Further, substantial retardation of this developmental phase of slow learning, even if caused entirely by environmental circumstances, is likely to prove almost irreversible. Firstly, in the absence of adequate stimulation and activity, neurophysiological mechanisms involved in normal adult learning may fail to develop (Haywood & Tapp, 1966; Hebb, 1949). Secondly, and perhaps more importantly, conditions in impoverished environments, such as custodial institutions, generally do not provide variety and duration of exposures to diverse elementary perceptual-motor experiences comparable to those available to the normal child during the first few years of life. Efforts, however persistent, to train or educate the older retarded individual by means of the teaching and training techniques appropriate to a normal individual of similar chronological age (CA) must be expected to prove largely ineffectual, since they do not compensate for the retarded individual's impoverished repertoire of stimulus and response patterns.

Selective Attention. Development of the ability to react to recurring stimulus compounds as unitary patterns and the integration of recurring response sequences into higher-order units, important though they are, would not suffice to account for the efficiency of the mature individual in learning, often on a one-trial basis, relations involving new combinations of properties or events as, for example, the names of newly encountered people or the identifications of new plants or animals. An important factor in the development of this ability is the acquisition of habits of selective attention with respect to the stimulus properties which differentiate the members of important classes of objects or events. Thus, what is termed "reading readiness" involves acquisition of the habit of distinguishing printed black-and-white patterns such as letters on the basis of contours, while ignoring differences in size of the retinal image or position in the visual field.

Beyond facilitating the learning of particular discriminations, the development of habits of selective attention for the stimulus properties which are relevant over whole classes of tasks or situations prepares the way for the emergence of organization in identification and recall. A relatively naive or immature individual, in learning to identify members of a class of objects, must begin with an essentially random process. On early trials random subsets of the possible associations between each object encountered and the proper identifying response may be formed. Since, in general, the objects in the class will share many stimulus properties or components, the associations thus formed will not suffice to permit identification without confusion. Thus, on subsequent trials the individual must sample additional elements, then rehearse and consolidate additional associations until enough have accumulated to provide a basis for reliable discrimination of all objects in the class.

At a later state of development, the individual may have learned by experience which stimulus dimensions are relevant and which are irrelevant for distinguishing members of a particular class of objects (Bower, 1967). Then upon encountering each new instance he need only note the "values" of the given object on a limited number of stimulus dimensions, and by utilizing associations involving these can immediately establish the basis for efficient identification. Thus, a child who has grown up in a western European culture, upon encountering a group of people from an Asiatic country, finds it very difficult to associate names with faces and complains that these people "all look alike." In this complaint, the child is expressing the fact that the members of the newly encountered group resemble one another on a number of the important stimulus dimensions (for example, hair and eye color) which he is accustomed to using as a basis for identification and recall.

Observational Learning. Although detailed experimental analyses are lacking, it seems clear on theoretical grounds that combinations of these developmental processes that have been separately identified may generate at least part of the higher-order learning phenomena which characterize the normal human being beyond the earliest developmental stages and which appear superficially to be qualitatively different from forms of learning observed in the young child, in the mentally retarded, and in animals. An important case in point is the ability of the normal human being to profit extensively from the experience of others. The efficiency with which older normal children and adults avoid most of the dangers in their environments cannot be explained in terms of escape and avoidance conditioning, for most of the potentially traumatic events are never experienced. Similarly, much constructive, goal-directed activity cannot be accounted for in terms of simple reinforcement principles, because the rewarding consequences toward which the behavior appears to be directed have never previously been experienced by the particular individual.

An important factor in this type of behavior is undoubtedly the learning of associations between behaviors and rewarding or punishing consequences by observation of other individuals, which Bandura (1965) terms "learning from models." Learning from models would be difficult to account for within the framework of the traditional law of effect or reinforcement theory, in which the basic conditions of behavioral modification are held to be selective strengthening or weakening, as a result of immediately following rewards or punishments, of responses made by the individual. Within the theory outlined in the preceding section, however, the role of rewards and punishments in learning is always indirect in the sense that the organism must first learn relationships between behavioral events and rewarding or punishing consequences, and then modify its process of response selection on the basis of anticipated outcomes.

The first stage in the development of the ability to learn from models must be the process of stimulus patterning we have already discussed. The organism must learn to recognize environmental events which are sources of reward or punishment regardless of whether these occur as a consequence of its own actions. Once this learning has occurred, the organism will be in a position to learn relationships between these events and others simply by perceiving them in temporal contiguity. Clearly this process can proceed without verbal mediation; for example, John, Chesler, Bartlett, and Victor (1968), gave a clear experimental demonstration of virtually instantaneous learning by mature cats to press a bar for food following the opportunity to observe other cats pressing the bar which activated the food-dispensing mechanism. The cats which apparently learned by observation did not, it should be emphasized, imitate the precise movements made by the "models"—rather they behaved as though they had learned to anticipate that movement of the bar would be followed by the appearance of food in the dispenser and thus, when permitted to do so, modified their own behavior by selecting responses which brought them into contact with the bar and resulted in its movement, even though these were not the same response patterns exhibited by the "models."

The researches of Bandura and his associates have begun to generate an account of the way in which this important form of higher-order learning develops in the normal child (Bandura, 1965; Bandura, Grusec, & Menlove, 1967). However, I have seen no reports of studies of this process in specific relation to mental retardation; if this area has indeed been overlooked in research to date, it seems almost certain that it would repay serious investigation. It seems almost inevitable that one of the respects in which the environment in custodial institutions must differ most drastically from that in normal homes and schools is with respect to opportunities for constructive learning from models.

Development of Motivational Systems. There is no reason to doubt that the motivational systems and emotional dispositions which contribute to normal adult behavior are the result of as prolonged and intricate a developmental process as learning and problem-solving abilities. However, experimental analysis lags, owing to greater difficulties of measurement and control. Consequently any interpretation of the development of motives in terms of learning theory must at present be somewhat speculative and fragmentary. It is easy enough to contrast the behavior of the very young child, closely bound to the effects of immediate reward and punishment, with the extended and organized behavior systems of the adult which are often only remotely and tenuously related to rewarding or punishing consequences. But at present it is difficult to do much toward filling in the gaps in the developmental sequence leading from the former to the latter. We do know that organized motivational and emotional systems do not emerge from an original condition of randomness. Even in the very young organism, reaction tendencies associated with various unconditioned stimuli show clear evidence of organization into systems. The taste of food by the hungry infant leads to facilitatory feedback which increases the probabilities of a considerable range of approach and ingestive behaviors, whereas a painful stimulus produces inhibition of positive systems and facilitation of the families of defense and withdrawal reactions. By virtue of the associative process discussed in the previous section, the events which precede these unconditioned rewarding and punishing stimuli become assimilated into the motivational systems; approach behaviors come to be facilitated by stimuli that have preceded food, withdrawal by stimuli which have preceded pain.

The balance of reward and punishment in an individual's environment is doubtless of the greatest moment with regard to the course of behavioral development. An ample supply of rewarding events must be available if the young organism's tendencies to explore its environment are to be strengthened and lead to the acquisition of a normal repertoire of behavioral sequences leading to various types of incentives and goals. Probably almost equally important for normal motivational development is an environment relatively free of punishment. According to the theory of punishment which seems to me best established (Estes, 1967, 1969) and which has some rather cogent neurophysiological as well as behavioral support (Miller, 1963; Stein, 1964), an important consequence of the traumatic or painful event is that it leads to immediate inhibition of currently active positive drive systems. As a result of conditioning, events which precede punishment would thus in turn come to produce inhibition of positive drive systems. Consequently, extensive experience with punishment would have the result of crippling the development of positive motivational systems. I hasten to add, however, that this observation should not be taken too broadly as a prescription against the use of punishment for controlling behavior of children, retarded or otherwise. The implication is, rather, that the use of punishment should always be combined with, or, to put it differently, compensated for, by reward for more desirable alternative responses. As, for example, in the work of Lovaas, Schaeffer, and Simmons (1965) with psychotic and autistic children, where even rather severe punishment may not have obvious effects in retarding the development of positive motivational systems if combined with lavish administrations of rewards.

On the basis of learning theory it must be assumed that the development of more generalized "achievement motives" (McClelland, Atkinson, Clark, & Lowell, 1953) depends upon occurrence of appropriate sequences of success experiences in earlier life, owing either to good fortune or to good management on the part of parents or educators. Thus, to take a relatively simple example, in the case of a boy who has had a sequence of successes in solving progressively more difficult arithmetic problems, with consequent indications of approval and esteem from teachers and parents (perhaps quite in contrast to the results of his efforts in other school subjects), the identification of new problems and the experience of intellectual effort in attempting a solution will come to be signals for anticipated reward. As the process continues, the behavior of seeking for new problems similarly becomes assimilated into the system and the child may be said to demonstrate motivation for achievement in this particular area. Conversely, as a result of repeated failure with a particular type of problem, the appearance of a new problem would constitute a cue for avoidance. The analysis of this type of motivational development in terms of the present learning theory does not suggest that failures need be entirely prevented in order to permit development of positive motivational systems, but rather that failures should be followed relatively promptly by successes on a variety of occasions so that the experience of failure followed by continued effort comes to function as a signal for anticipated reward rather than for continued failure.

Internalization. A major overall trend in the development of normal modes of adult learning and performance is the progressive internalization of behavioral processes which are originally overt. At a relatively simple level, this trend is observed in the development of discrimination-learning ability, even in animals. The simplest discrimination problems involving differential response to stimuli which have elements in common cannot be solved if the organism responds overtly to a given stimulus display without first inspecting all of the elements or component features. If an overt response is evoked by the first cues sampled, sometimes these will be the elements common to the stimuli being discriminated. Over a series of trials in almost any discrimination task premature responses lead to decreases in frequency of reward, and the organism learns to slow down at choice points and expose itself to the relevant cues before making an overt choice.

In the normally developed human being, this type of learning generalizes to new situations. Ultimately the individual tends, when any new problem situation presents itself, to withhold overt response until the alternative response sequences which might be initiated have been examined covertly and the one selected which generates the highest expectation of a rewarding outcome. This process of covert scanning of alternatives prior to choice is evidently conspicuously lacking in the mentally retarded, who are commonly characterized by such terms as "impulsive" and "distractible." The sources of this type of deficit are, however, little understood and doubtless merit further investigation.

It seems to be a rather common assumption that there may be some relatively localized organic basis for the inability of the retarded individual to inhibit premature responses, or to resist distraction by irrelevant stimuli during the period of choice. However, one cannot at present rule out the possibility that these deficits are largely a result of an absence of the necessary conditions for the learning of habits of stimulus scanning and "vicarious trial and error." On theoretical grounds we cannot expect habits of this type to generalize to broad classes of situations until they have first been learned in relatively specific contexts. And meeting this condition must require experience with sequences of problems of appropriately graded difficulty which permit individuals to learn by experience that some modes of responding lead to substantial increases in the likelihood of reward. If discriminations encountered during the early developmental stage are typically too difficult, so that attempts at solution lead to no detectable increases in reward input, we must expect that the general response tendencies learned will be those which most quickly terminate exposure and lead to escape from the frustrating situation.

Part B

Empirically Oriented Research

IV. *Heritability and Development of Learning Ability*

My main purpose in this study is to analyze the role of learning theory in research on mental development, not to survey the empirical literature. In this section I have attempted only to characterize the empirical side of the interaction between theory and experiment, the relationship which is the main focus of attention in the remainder of the volume. From the very large body of research reports considered, I have selected those which seem directly relevant to this relationship. Where appropriate I have discussed the adequacy of methodology and the interpretability of findings in terms of the presently available body of learning theory summarized in Chapter I. The empirical findings cited in this chapter will give the reader some perspective with respect to the body of experimental facts presently available as a basis for theoretical development.

1. genetic studies

In view of the well-documented contribution of inheritance to individual differences in intelligence (see Hilgard & Atkinson, 1967, for an introduction to the extensive literature), surely few psychologists doubt that variation in learning ability, within as well as between species, is in part genetically determined. However, positive support for this belief is based largely upon general biological principles rather than upon specific research findings. The relatively new discipline of behavior genetics has advanced rapidly during the last two decades (Fuller & Thompson, 1960; McClearn & Meredith, 1966) but still appears considerably short of being able to shed light on the heritability of specific learning capacities.

Much of the earlier research in this area has gone into demonstrations that strains of animals which differ markedly with regard to performance in standard learning situations can be generated by selective breeding. Not surprisingly, work relating behavioral differences to chromosomal variables has, so far, been limited to studies of lower organisms, principally *Drosophila* (Hirsch & Erlenmeyer-Kimling, 1967). Large differences between strains of rats in rate of maze learning (Heron, 1935, 1941; Tryon, 1940) or between strains of dogs in rate of conditioning (Fuller, 1955) are readily produced by selective breeding, but scarcely a start has been made toward analyzing the behavioral differences between strains in terms of specific learning or motivational processes.

The first attempts to relate differences between maze-bright and maze-dull strains of rats to specific, theoretically significant aspects of learning were not very fruitful, in part because of the primitive state of the associated genetic analysis (Estes, 1942; Heron & Skinner, 1940). Recent work with highly inbred strains of mice appears rather more promising. For example, Bovet, Bovet-Nitti, and Oliverio (1969) report striking differences between strains with respect to effects of temporal variables upon avoidance conditioning and retention of avoidance responses and suggest an interpretation in terms of a two-stage conception of memory storage. Although one still can find few points of direct interaction between learning theory and behavioral genetics, there are grounds for expecting a more fruitful relationship between these disciplines to develop.

2. developmental studies

Presumably information relevant to the maturation of elementary learning capacities must come largely from analyses of simple forms of learning in children at the earliest stages of development, perhaps primarily during the first year of life. However, research on learning during this period is beset with difficulties and most of the effort to date has gone into developing appropriate methods.

Even the fragmentary available evidence, based on experiments conducted with differing procedures at different age levels, leaves little doubt about the gross empirical fact that conditioning is very difficult to demonstrate during the first few weeks of life. Studies of classical aversive conditioning in infants 1-3 weeks old have yielded conflicting and tenuous results, and " . . . it must certainly be concluded that such conditioning is at best difficult to obtain" (Lipsitt, 1963, p. 162). Evidence for classical appetitive conditioning in infants of this age range is somewhat firmer, positive results having been reported by Irwin (1930), Marquis (1931), and Lipsitt and Kaye (reported in Lipsitt, 1963), among others. Simple operant conditioning has been demonstrated in infants 3-4 months of age by Brackbill (1958) and Weisberg (1963), learning being manifest in the former case by "returning a smile" and in the latter by changes in vocalization as a result of "social" reinforcement. An operant discrimination between red and blue stimuli by 12-month-old infants has been reported by Simmons (1964). In no case do there seem to be available any data based on fully comparable conditioning experiments conducted on infants at a sequence of age levels. A review by Lipsitt (1963) indicates that, although empirical results consist largely in demonstrations that conditioning and other simple forms of learning can be produced in infants, advances in methodology are increasing the prospects for more analytical work.

Even at the animal level there have been few systematic studies permitting well-controlled comparisons of learning functions on the same tasks at different ages. Reviewing a series of studies of subhuman primates, Harlow (1959a) reports that for monkeys, classical conditioning was as rapid at 5 days of age as in adulthood. Rate of object-discrimination learning, on the other hand, increased substantially with age. Trials to criterion decreased from about 200 at 30 days to an asymptotic level of about 10 at 150 days in comparisons involving animals first exposed to the experimental situation at the given age. The younger animals persisted longer than adults in object or position preferences and shifted more often from one to another incorrect hypothesis.

It would be of prime interest from a comparative standpoint to determine whether similar trends obtain for human children. The most relevant study that has come to my attention was reported by Hill (1965). Groups of children were first tested on object-discrimination, oddity, and conditional-oddity learning at ages 1, 4, 6, and 12 years. The "50% thresholds," that is, the ages at which 50% of children tested were able to meet a learning criterion, were found to be at 1 year for object discrimination, at 6 years for oddity problems, and 6-12 years for conditional oddity. At all ages, children mastered any given discrimination within about 40 trials if at all. Backward learning curves were similar in form to those reported by Zeaman and House (1963) for discrimination learning by retardates; initial plateaus at a chance level of responding terminated in abrupt transitions to criterion performance. The hypothesis which immediately comes to mind is that children whose performance curves did not depart from the initial plateau on any given type of problem were, like the immature monkeys, not responding randomly but rather persisting in position or stimulus preferences which were incompatible with solution.

A bit of information possibly bearing on this hypothesis is to be found in a developmental study of color-form discrimination reported by Milgram and Furth (1964). Groups of 6- and 9-year-old normal children and retarded groups matched to them for MA were given training on original discriminations followed by various shifts, of which the positional reversals and dimensional reversals are of particular interest. During original learning, the correct cue on the relevant dimension, either color or form, always appeared in the same position. On positional reversal, the color or form dimension became irrelevant and the correct position was changed. On dimensional reversal, the position assignment remained the same, but correct and incorrect cues on the relevant dimension were interchanged, so that performance should be susceptible to substantial facilitation by attentional or mediating responses. In terms of errors to criterion, clear relationships to age and normal-retarded differences within age groups were exhibited only on dimensional reversal.

Similarly, in the case of probability learning, the progressive increase with age in rate of approach to a probability-matching level of performance has been shown to be attributable at least in part to the greater tendency of younger children to persist in systematic, but inappropriate response patterns, such as alternation (Craig & Myers, 1963; Stevenson & Odom, 1964). In some variants of the probability-learning experiment, however, alternation tendencies have been found to increase with age to a maximum and then decrease, suggesting that the shifts in response patterns may reflect the learning of habits of approach to specific types of problems rather than a simple maturational process (Weir, 1964).

In view of the considerable body of theory now available to interpret experiments on the simpler forms of verbal learning, it seems that well-controlled comparisons of, for example, paired-associate learning, free verbal recall, and concept identification at different ages, might substantially further our understanding of the development of verbal learning abilities. Unfortunately, reviews of the literature on verbal learning in children reveal a poverty of relevant studies (Goulet, 1968; Keppel, 1964).

I have been able to find data for the same experimental paradigm over a wide range of age levels only in the case of free recall. Amount recalled from a given list has been found to increase with age from 5 years to college level, and this function is roughly paralleled by that for a measure of clustering in recall (Bousfield, Esterson, & Whitmarsh, 1958; Cole, Frankel, & Sharp, 1970, in press). Further, the tendency to organize recall by means of clustering has been observed to vary but little over trials within an experiment in the case of 5-10-year-old children but to increase steeply in the case of college students and older adults (Laurence, 1966).

A recent study by Moely, Olson, Hawles, and Flavell (1969) yields evidence that the child's ability to utilize clustering when appropriately instructed and his tendency to cluster spontaneously follow quite different developmental trends. Children in kindergarten and first grade exhibited little spontaneous clustering and increased their clustering activity only in response to specific instructions, whereas older children required little prompting. However, at all ages clustering, once initiated, improved recall. Cole, Frankel, and Sharp (1970) report that increases in recall as functions both of age and of amount of clustering are largely confined to the early and middle portions of the recall list, as would be expected if the increases reflect more effective strategies for retrieval from long-term memory. Evidently the spontaneous use of rather general strategies for improving retention tends to increase as a function of age, but little is yet known as to the specific kinds of experiences which are responsible. A promising direction for further research is suggested by the success of Kingsley and Hagen (1969) in producing shifts toward the adult pattern when they gave young children specific training on retrieval skills for their serial-order memory task.

Age-related trends in retention of nonverbal material may be of more relevance for the analysis of maturational processes, but even here it is difficult to partial out the role of verbal mediation. To take one example, Paraskevopoulos (1968a, b) has conducted a rather extensive developmental study of short-term memory for dot patterns in retarded, normal, and gifted children. Age groups ranged from 6-7 to 11-12 years for the gifted and from 8-9 to 15-16 years for the retarded. Each group was tested on random dot patterns and on patterns characterized by bilateral, horizontal, or double symmetry. For each condition, retention scores, in terms of mean errors in reconstructing patterns from memory, declined as an orderly function of age. The pattern of relationships among the separate curves for patterns with different types of symmetry was very similar for the retarded and gifted groups, but the gifted's rates of decline of the age functions were greater and the apparent asymptotes lower across the board. It is difficult to judge, however, the extent to which these trends can be taken to support the author's interpretation in terms of development of perceptual organization, for patterns characterized by different types of symmetry are amenable in differing degrees to verbal recoding and rehearsal.

More nearly direct evidence concerning the operation of linguistic habits in producing age-related changes in retention comes from a study by Flavell, Beach, and Chinsky (1966). Records of children's spontaneous verbalizations showed an increase from age 6 to age 11 in likelihood of rehearsing stimulus names in a nonverbal serial recall test. Further, within the middle age group (the only one for which the breakdown was feasible) verbalization was directly related to retention on both immediate and delayed recall tests.

Theoretically relevant data concerning developmental changes in learning processes have been forthcoming largely for the period from ages 4-7, the first few years of school. The relatively heavy concentration of effort upon this period is understandable since it is during these years of childhood that the individual's behavior comes to a marked extent under the control of language and therefore that modes of learning in the child begin to depart markedly from those found in the lower animals.

One of the earliest major studies of discrimination learning by children in this age range demonstrated an orderly increase in the tendency to transpose stimulus relationships (Kuenne, 1946). In another well-known series of studies, the Kendlers (Kendler and Kendler, 1962; Kendler, 1962) found a change during this period from faster nonreversal shifts in younger children to faster reversal shifts in older children following similar original discrimination training. Also, they found a progressive increase over the age range 3-10 years in the proportion of children who reversed a previously learned discrimination when the second task permitted a choice between a reversal and a nonreversal shift.

These types of findings have been taken to support the idea that within this age range the child comes increasingly to verbalize stimulus relationships and that verbal mediating responses facilitate the transfer of discrimination learning to new situations. A similar interpretation may be applicable to White's (1966) findings that the effects of varying the positive stimulus, the negative stimulus, or the relative positions of positive and negative stimuli upon rate of discrimination learning shift from negative to neutral or positive during the age range of about 5-7 years.

Summarizing the research literature on learning in children between 5 and 7 years of age, White finds considerable evidence for three rather general developmental trends: (*a*) increasing use of linguistic mediating responses, labeling, categorizing, and the like, as aids to learning and performance; (*b*) increasing ability to maintain orientation toward "invariant dimensions of stimuli"; (*c*) increasing ability to "string together representations of stimulus-response-consequence into sequences which, projected into the future, allow planning, and projected into the past, allow inference" (1965a, p. 210).

On the whole, it appears that learning in children prior to about age 5 proceeds in similar fashion to the learning observed in comparable experiments with animals. Except perhaps for some aspects of linguistic development, this early learning is directly interpretable in terms of associative principles. But, although the character of human learning changes notably beyond age 6 or 7, with steeply increasing control of behavior by verbal rules and strategies, it need not be assumed that associative processes cease to operate.

White (1965 a, b) has assembled a number of lines of empirical support for the conception that a hierarchical organization of learning processes develops with age. The older child, as the adult, continues to learn in accord with associative principles, but also comes to inhibit responses arising directly from stimulus-response associations when these conflict with higher-order, cognitive processes. A novel source of evidence which White (1965b) brings to bear upon this hypothesis is the relationship between correctness of response and time taken to respond in discrimination learning. For the youngest children tested (age 3½-4½) the lowest latency responses are most likely to be correct, and the curve of proportion correct versus latency increases monotonically, being apparently asymptotic beyond 2 sec. With increasing age, this function takes on an inverted S shape. For older children (age 5½-6½), and for adults on stimulus generalization tests, probability that a response is correct decreases as a function of latency to a minimum at .25-.50 sec, increases to a maximum at .80 to 1.0 sec, then decreases at longer latencies.

White interprets this latter function in terms of a "temporal stacking" of alternative responses which may be initiated by a given stimulus. In the course of learning, the older child or adult comes to inhibit his initial response tendencies to stimulus presentations and performs in accord with verbal rules or other higher-order modes of response. On some occasions, especially under stress, the higher-order processes become temporarily inoperative, leading to the disinhibition of incorrect stimulus-response associations.

V. *Empirical Comparisons between Normal and Retarded Populations*

The vast majority of studies of learning in the retarded up to the present, and virtually all of those done before the first major reviews of the literature appeared in the early 1960s (e.g., Denny. 1964), were concerned with comparisons between groups of normal and retarded subjects in various standard laboratory situations. The groups undergoing comparisons have ordinarily been defined by performance on standard intelligence tests, but occasionally by grosser criteria such as success or failure in school, or ability to get along outside a custodial institution. With the exception of a few current research programs, which will be discussed in detail in later sections, these researches have not been guided by analyses of the problems of retardation in terms of learning theories. The underlying assumption, never fully explicated, seems to have been that these group comparisons might somehow aid in localizing the sources or types of deficit in intellectual functioning that characterize the mentally retarded.

1. methodology

Within the framework of comtemporary learning theory, how can we characterize what is being done when intellectual ability is appraised by means of standardized tests? Samples of performance are taken on a variety of tasks, many of which would be expected to involve learning explicitly (for example, tests of general information, vocabulary, arithmetic reasoning) but usually with some admixture of tasks that are assumed to be relatively free of effects of specific previous learning experiences. Since the previous learning relevant to test performance has gone on outside the testing situation, we are sampling performance at particular points on learning curves, generally with no picture of the form of the learning function and with no way of identifying the points on the abscissas at which samples are being drawn. We know that in virtually all of the learning situations that have been studied in the laboratory, individual differences appear in initial levels of performance, terminal levels of performance, and rates of approach to asymptotic levels (see Gagné, 1967, for reviews of the literature relevant to this generalization). This ubiquitous variation in parameters of empirical learning functions is almost universally assumed to be partly attributable to previous experience and partly to organismic variables which may in part reflect innate structural differences among individuals.

Why should test scores based on these samples of behavior in very limited situations predict performance in school and in everyday life? The methods of test construction insure this result by arranging that, to a considerable extent, the tests sample what the individual has learned up to the age of testing. To the extent that the sampling is adequate, an individual's standing in a group with respect to the skills and knowledge represented in the test must, statistically, predict his relative standing with respect to the larger population of skills and knowledge being sampled. Further, if different individuals tested at a given age have had the same opportunities to learn, test performance must in part reflect parameters representing rates of gain, that is, the extent to which these individuals have profited from the same amount of previous experience, and thus should predict rates of gain from experience in new situations. If, however, opportunities to learn differ among these individuals, then, of those receiving the same test scores at a given age, the ones having had fewer previous opportunities to learn must have higher rates of gain and thus should be expected to gain more from the same amount of experience on new tasks.

Within this framework, what is being done when normal and retarded individuals are compared on specific laboratory learning measures? What could one hope to accomplish from such comparisons? What controls should be employed? When we define groups of normal and retarded learners on the basis of test scores and compare these groups on a variety of laboratory tasks, we obtain an alternative description of differences in performance between these groups. What might we hope to gain by replacing a description of differences between the groups in terms of subtests of the Stanford - Binet scale or the like with a description in terms of differences in performance in eyelid conditioning, maze learning, paired-associate acquisition, and similar laboratory learning tasks?

It seems likely that one main source of motivation for these studies is similar to that of the investigators of mental abilities who have attempted to reduce the concept of general intelligence to one of variation in a small number of traits or factors (e.g., Spearman, 1927; Thurstone, 1938). The analogous objective here might be to reduce the description of differences in mental ability now given by scores on heterogeneous tests to description in terms of differences revealed by laboratory experiments in a small number of presumably more unitary and basic functions or capacities (for instance, conditionability, susceptibility to inhibition, decay rate of stimulus traces).

Even if the preceding, relatively ambitious objective should prove unattainable, one might hope to replace the description of mental ability in terms of a broad cross section of performance on various tasks for which we have little knowledge of boundary conditions with a description in terms of performance and rates of gain in performance in certain standard situations which are better understood and more amenable to manipulation and analysis. To anticipate the conclusions of this chapter, I should be inclined to suggest that progress in this second respect might represent the principal fruit to be anticipated from empirical comparisons between retarded and normal individuals in laboratory situations.

With regard to these objectives, a number of important methodological points need consideration. Firstly, we may envisage the possibility of getting directly at rates of gain in performance rather than having to infer these from cross-sectional measures of performance under the assumption of equal opportunity to learn up to the time of testing. There are as yet almost no relevant data for this purpose because durations of laboratory learning studies typically are substantially shorter than the time periods involved in school learning or the development of the kinds of skills and knowledge reflected in tests of mental ability. Recent developments in computer technology and its applications to teaching and training may make it possible within the relatively near future to conduct longer-term learning experiments of the kind needed to provide assessments of rates of learning over intervals of weeks or months as opposed to the single brief sessions involved in most learning studies to date.

Secondly, major attention is needed to the problem of controlling adequately for opportunities to learn the material involved in assessments of mental development. Differences in learning ability (and therefore in the rates of improvement to be expected on new tasks) can be inferred from performance at a given time only if previous opportunities to learn have been equated for individuals being compared. It must be emphasized further that control is required not only for previous opportunities to learn the specific material involved in the given test, but also for previous opportunities to learn other habits or skills or to develop motives or emotional dispositions which might influence performance at the time of testing.

Lack of regard for this point in the literature on mental retardation is hard for the outsider to credit. It is not uncommon to see comparisons of performance on laboratory tasks between groups differing as radically with regard to previous opportunities to learn as institutionalized retardates and college students of similar chronological age. Under such circumstances differences in performance on the reference task must be expected to reflect some unknown mixture of effects of differences in past opportunities to learn and in the type of subject variable associated with the notion of capacity (not to speak of other determiners of performance such as motivation and specific skills in dealing with test situations).

On theoretical grounds I can see no way to avoid, and little room to ameliorate, the conclusion that when differences in previous opportunities to learn and the relationships of these to the reference task are unknown, comparisons between different groups of individuals are of little value for any theoretical purpose. Investigators in this area would doubtless reply that it is difficult in practice to equate groups of subjects on variables other than test scores or chronological age. This difficulty is surely amply appreciated by all who have had any familiarity with this research area. Nonetheless, the fact that better controls will be difficult to achieve adds nothing to the value of uninterpretable data. I am inclined to press this point strongly, for it seems generally to be glossed over, perhaps out of sympathy with the many difficulties of working with retarded groups.

Granted that the ultimate desideratum of perfect matching of groups on the basis of previous opportunities to learn is impossible to achieve, much more can and must be achieved with regard to approximating it. Even at the cost of reducing the number of experiments being conducted, those which are done with a view to comparing normal and retarded populations could be expanded to include multiple normal control groups differing in the degree to which they match the retarded group with respect to variables related to previous learning opportunities. By this means a beginning could be made toward separating the contributions of opportunities to learn and individual differences in capacity with respect to learning situations in which these sources of variation are now inextricably confounded.

A third principal methodological question concerns the selection of normal control groups to be compared with a given group of retarded subjects. Comparison groups equated on CA and on MA appear somewhat unsystematically in the literature. To my knowledge, the issues involved were first discussed explicitly by Denny (1964). Upon theoretical analysis, the following considerations emerge. If previous opportunities to learn are in fact equated for the groups of subjects being compared, then among groups having the same CA, those with higher IQs will differ with respect to parameters having to do with rate of learning as well as with respect to present levels of the skills and knowledge sampled on the test. This is so, since the higher-IQ groups would have profited more from equal opportunities to learn over the same period of time. One might expect, then, by means of these comparisons, to localize the differences in learning rates with respect, for example, to particular types of tasks or intellectual functions, and to investigate conditions that might modify the differential learning rates in laboratory learning situations.

If, given the same premise, different groups are equated with respect to MA scores, thus differing with respect to CA, the higher IQ, and therefore younger, groups will have gained more from the same opportunities to learn and thus must rank higher on capacity variables. Therefore, of the groups matched on MA, those higher in IQ would be predicted to gain more from new experiences, presumably including opportunities to learn in laboratory situations. However, in attempting to pursue this strategy, one arrives at a logical impasse, for previous opportunities to learn cannot really be equated even in principle if age at testing differs. Thus there seems to be no rational basis for expecting progress toward isolating differences in capacities by means of comparisons between groups matched on MA—at least until very much more is known as regards how to equate for opportunities to learn over different time spans.

Even the conclusions sketched above regarding groups matched on CA must be weakened greatly if one cannot assume that previous opportunities to learn are also equated. For under this, probably all but ubiquitous, circumstance, the performances of groups equated on CA but differing in IQ will reflect some mixture of differences in previous opportunity to learn and differences in capacity parameters. One of the few types of comparison apt to be satisfactorily interpretable will be that involving tasks unlikely to be influenced by previous learning histories to any appreciable extent (for example, classical conditioning).

Thus comparisons of learning rates and other aspects of performance in classical conditioning and other simplified and "artificial" learning situations which differ radically from those to which subjects would have been exposed previously in everyday life would appear to provide the most promising starting point for analyzing differences in capacity variables between normal and retarded populations.

Again, comparisons between groups matched on MA but differing in IQ would be still harder to interpret. The high-IQ individuals would have reached the common MA at an earlier CA owing to some mixture of higher capacities and richer opportunities to learn. When the former aspect is dominant, the higher-IQ subjects of a given MA should do better on new tasks, since they would tend to have the higher learning rates. But when the latter aspect is dominant, the lower-IQ subjects, having reached the same level of performance with fewer opportunities to learn, must have the higher learning rates and should be expected to do better on new tasks. Thus, at the present stage of knowledge in this area, equal MA comparisons would seem to be almost uninterpretable in the vast majority of learning situations. The obvious justification for using both equal CA and equal MA comparison groups is that the pattern of differences on different kinds of tasks may shed some light on the probable roles of previous experience versus capacity parameters—and this much is apt to be achieved only if more attention is given to varying the composition of comparison groups systematically with respect to average standing on variables related to previous opportunity to learn.

A third possible fruit of normal-retardate comparisons in the learning laboratory is the isolation of differences between these populations on variables which are significant within well-developed theories, thus promoting the interpretation of the differences in biological or neurophysiological terms on the one hand, and facilitating the search for means of preventing or modifying retarded mental development on the other. Almost by definition, progress in this direction is unlikely to arise from empirically directed, comparative studies. Thus it is not surprising that almost no signs of such progress are to be seen in major reviews of the empirical literature up to the early 1960s dealing with learning in the mentally deficient (Denny, 1964; Lipman, 1963; Stevenson, 1963). Later sections of the present review will deal with recent, somewhat more promising, trends appearing in research programs involving more theoretical direction.

2. comparative experimental studies

With these methodological considerations in mind, it may be useful now to review briefly the principal findings that have emerged to date from laboratory studies designed primarily to compare learning in normal and retarded groups. A detailed survey of individual experiments and results seems unnecessary, since this has been largely accomplished by Denny's 1964 article together with a number of others appearing in the first three volumes of the *International Review of Research in Mental Retardation* (Ellis, 1966a, 1966b, 1968). Thus, I shall, in effect, take Denny's very thorough and critical analysis of the literature up to 1964 as a starting point and extend or modify his conclusions as indicated by more recent work.

It would be pleasant to report that the very large acceleration of research on mental retardation during the past five years has made obsolete the picture of the state of knowledge in this area emerging from Denny's critique. Unfortunately, the contrary appears to be the case. With few exceptions the general methods and assumptions underlying empirical comparisons of normal and retarded learners in laboratory tasks have changed relatively little, and perhaps because of this inertia with respect to innovations in methods and outlook, a diminishing-returns effect seems clearly manifest.

Classical Conditioning. The literature on classical conditioning in the retarded comprises three main components: a few early studies by American investigators dealing with either salivary or eyelid conditioning (Denny, 1964); a considerable body of Russian literature, the relatively small portion available in English having been reviewed by Denny (1964) and Astrup, Sersen, and Wortis (1967); and a small number of recent studies by American and British investigators, largely utilizing eyelid conditioning, which will be discussed in this section.

Perhaps the most striking single finding running through the conditioning studies is the absence of any conspicuous difference between acquisition rates of normal and retarded groups, regardless of whether the comparison is made with CA or MA equated, or of any substantial correlation between acquisition rate and IQ within retarded groups (Cromwell, Palk, & Foshee, 1961; Doubros, 1965; Franks & Franks, 1962; Grings, Lockhart, & Dameron, 1962). In some cases slight advantages have been observed for normals when appetitive USs (for example, food as a stimulus for salivation) have been used and a slight reversal of this trend when aversive stimuli have been used. Luria (1963) reports approximately equal rates of conditioning for "child oligophrenics" and normal peers when a verbal stimulus was used in the role of a US.

In extinction, on the contrary, though the data are by no means uniform, the preponderance of studies show faster extinction in normals than in retarded subjects of similar CA. Primarily on the basis of this frequently observed extinction differential, it has been suggested (e.g., Denny, 1964) that the retarded learner is characterized by a pervasive inhibition deficit. This notion is intuitively appealing in some respects. Though without much specific documentation, there is a general tendency to associate mental retardation with inferior cortical functioning, and there is considerable reason to believe that an important function of the cortex in conditioning must be the selective inhibition of responses and response components other than the precise CR undergoing acquisition. Also, it is easy to imagine extensions of the notion of inhibition deficit to aid in accounting for inferior performance by retardates in such situations as delayed reaction and discrimination learning. Further, an inhibition deficit on the part of retardates might account for the findings of Ross, Koski, and Yager (1964) to the effect that, within a retardate group, acquisition curves for eyelid conditioning under 100% and 50% reinforcement schedule were virtually identical and in sharp contrast to the direct relation between percentage of reinforcement and rate of acquisition for college students.

Nonetheless, it must be pointed out that the concept of inhibition deficit is at present little more than a speculation with some slight empirical support. A difference in extinction rates need not reflect a difference in an inhibition process in any technical sense of the term *inhibition*. To point out just one quite different process which may be involved in these results to an important extent, a considerable body of evidence indicates that a CR may become associated during training with contextual stimuli, including traces of stimuli occurring on previous trials, as well as with the CS itself. When either human or animal subjects are given repeated sessions of conditioning and extinction it is generally observed that the rate of extinction becomes progressively more rapid. The interpretation usually offered is that the stimulus context changes greatly from acquisition to extinction, since during acquisition CRs are evoked in the presence of contextual cues from previous trials, all of which have terminated with USs, whereas during extinction the CS is presented on any trial in a sequence of previous trials which have not ended in USs. Differential tendencies to sample (attend to) these contextual cues could be responsible for the difference in extinction rate between normal and retarded groups that has been attributed to an inhibition deficit. The same factor could account for the difference reported by Ross *et al*. (1964) between college students and retardates in the sensitivity of acquisition functions to intermittency of reinforcement. Because of the potential theoretical importance of the concept of inhibition deficit, it seems essential that intensive analysis be continued in order to narrow down the range of alternative interpretations of the observed differences in extinction rate between normals and retardates.

Because it is one of the simpler and better-understood varieties of learning, and because the types of conditioning studied in the laboratory are likely to be relatively unaffected by specific transfer of training from previous learning experiences, classical conditioning would seem to offer a number of advantages as a basic situation in which to begin the analysis of relations between elementary learning processes and mental development. Ross (1966) points out further advantages, including the degree of control which can be achieved in conditioning experiments (bringing them closer than most other learning situations to the possibility of interpretation in terms of neurophysiological processes), the advantages for comparative and developmental studies arising from the fact that very similar experimental paradigms can be used with a wide variety of organisms, and finally the theoretical possibility that conditioning processes may underlie more complex forms of learning.

The negative side of the coin is that today very little progress can be claimed toward exploiting these advantages. In principle, conditioning experiments can be so arranged as to reduce drastically the roles of such variables as instructional and motivating conditions and to minimize or eliminate the processes, largely verbal, associated with the notion of voluntary control of behavior. Thereby, a start could be made toward analyzing differences between normal and retarded learners with respect to measures relatively unaffected by differences in previous learning histories. In studies reported to date, however, little appears to have been done toward achieving these simplifications and the conditioning performance of normal and retarded subjects must be assumed to reflect mixtures of learning and performance variables, the latter including such factors as differences in motives or attitudes, understanding of instructions, verbal expectations, and responsiveness to contextual cues associated with preceding trials of an experimental series.

Operant Conditioning. Although some learning theorists assume that operant conditioning is basically similar to, or reducible to, classical conditioning (Asratyan, 1967; Estes, 1959; Miller, 1959), the two learning situations are clearly distinguishable by the temporal relationships and the contingencies between the experimenter-controlled variables. Whereas in classical conditioning, training is given in discrete trials, each beginning with the CS and terminating with the US, in operant-conditioning experiments the organism is ordinarily permitted to respond at its own rate with reward contingent upon specified responses. Many studies of the retarded have been conducted using operant-conditioning methods, but the objectives have been characteristically quite different from those of classical-conditioning experiments. Whereas almost without exception the latter have been concerned with comparing normal and retarded individuals, the former have nearly always been attempts to see how far the behavior of relatively severely retarded individuals can be brought under the control of the same variables which are known to control operant conditioning and discrimination learning in normal human beings and animals. On the whole, a review of this literature gives the impression that even severely retarded individuals respond in qualitatively much the same way to various reward schedules and contingencies as do animals and normal human beings (see Denny, 1964, and Krasner & Ullmann, 1965 for summaries). However, special procedures may be needed to establish effective reinforcers for the severely retarded (Orlando, 1965). Detailed information as to how rates of acquisition and extinction of simple operant responses might be related to intellectual level seems to be entirely lacking.

Classical and operant-conditioning methods have offered special advantages for the investigations of learning in the severely retarded, with whom communication by means of instructions is difficult or impossible. The most recent review of this area, by Spradlin and Girardeau (1966), indicates that basic data upon responsiveness of the severely retarded to standard USs and CSs are still sparse, with no clearly established systematic differences between retarded and normal subjects. Operant-conditioning methods have been effective in training simple responses in even the severely retarded, using food, tokens, and some forms of social reinforcement. Though the data are by no means as orderly as those frequently obtained with animals, effects of different partial reinforcement schedules and discriminative contingencies appear generally similar to those established in the copious literature on operant conditioning of rats, pigeons, monkeys, and normal children (Bijou, 1957, 1958, 1961; Bijou & Orlando, 1961; Ellis, Barnett, & Pryer, 1960; Johnson, 1966; Oatley, Bryant, & Tinson, 1965; Orlando & Bijou, 1960, 1967).

Operant conditioning techniques are directly translatable into methods for the training of simple habits; thus in this area there is no sharp line between theoretically oriented research and the development of practical training methods. Numerous successes have been reported in the training of simple habits such as self-feeding, cleaning, dressing, and the like in the institutionalized retarded. For example, Birnbrauer and Lawler (1964) and Birnbrauer, Bijou, Wolf, and Kidder (1965) report substantial progress in training moderately and severely retarded children in numerous classroom routines such as hanging up coats, taking seats, working on programmed "prereading" lessons. Girardeau and Spradlin (1964) have begun a long-term project on the training of moderately retarded children to live in an institutional community. In the first stages tokens were established as rewards by permitting them to be exchanged for food and the like. Then training routines were developed to shape the children's behavior toward the establishment of such habits as caring for pets, making beds, group play, and simple work routines. Formal data seem not yet to have been reported, but the investigators report that "gross observations indicate that the desired behavior is increasing in frequency" (Spradlin & Girardeau, 1966, p. 288).

Perceptual-Motor Learning. Although simple perceptual-motor tasks appear to overlap little in content with those involved in standard intelligence tests, there is little reason to believe that the considerable literature comparing normal and retarded populations on perceptual-motor performance is likely to be of great analytic value. In general, performance on such standard perceptual-motor tasks as the pursuit rotor, digit-symbol substitution, and mirror drawing improves with age and must be assumed to depend on physical development and previous experience, neither of which have been controlled in comparisons of groups differing with respect to IQ. It has been found, for example, that retarded groups may do significantly better than equal MA normals on simple sensorimotor tasks, even though the same groups yield an advantage for the normals in rote verbal learning (Cruickshank & Blake, 1957).

The principal overall trends discernible in the literature (Denny, 1964; Lipman, 1963) are that on very simple tasks retarded individuals perform more poorly than normals of equal CA initially but tend to catch up with practice; on more difficult tasks the pattern is similar, the retarded gaining somewhat on the normals but normal groups appearing to approach higher terminal levels.

On the much-studied pursuit-rotor task, performance clearly is directly related to MA (Ellis & Sloan, 1957). Normal groups exhibit more rapid acquisition of skill within a single learning session and also exhibit larger increases in performance ("reminiscence") over rest intervals (Ellis, Pryer, & Barnett, 1960; Jones & Ellis, 1962).

One might be tempted to take the latter finding as a reflection of the same process underlying the presumed "inhibition deficit" in classical conditioning, but the analogy seems most tenuous. With relatively massed training, performance on simple motor tasks such as rotary pursuit typically levels off considerably below maximum attainable performance, and, in fact, it is not unusual to observe a decrease in performance late in the session. These effects have usually been attributed by theorists in the area of motor learning to the buildup of some kind of inhibition during the session. In terms of this concept, the observed difference might reflect differences in the amount of inhibition developed during the session or in rate of decay, but whereas the assumption of a faster rate of decay of inhibition for normals might seem to account in a simple way for the pursuit data, it would not be in harmony with the interpretation of extinction differences in terms of an inhibition deficit, for faster decay of inhibition on the part of normals would lead to more spontaneous recovery and therefore a slower, rather than more rapid, observed extinction rate.

In any event, before this apparent paradox could be taken very seriously, one would need to rule out experimentally such alternative interpretations as differences in covert rehearsal between experimental situations or differences in tendencies to develop and retain verbalizable tactics or strategies that facilitate performance.

In the learning of finger mazes or stylus mazes, which generally demand little skill in motor performance, earlier studies utilizing open mazes generally exhibited no appreciable difference in learning rates between retarded and normal subjects. When, however, the maze was shielded from view of the subject, a direct relationship between speed of acquisition and IQ emerged (Ellis, Pryer, Distefano, & Pryer, 1959). Since the learning of a shielded maze is clearly facilitated by verbalization of a sequence of turns, this latter finding may well be simply another example of the often-observed tendency for the retarded to be much less likely than normal individuals to develop and utilize verbal rules and strategies.

There seems no reason to doubt that capacity variables are involved in the determination of simple perceptual-motor performance, but neither does there seem to be a positive basis for assuming any appreciable correlation between these and the capacities involved in the more complex, and largely verbal, behaviors which are the basis for classifying normal and retarded populations. Although it is by no means certain that the differences between normal and retarded groups ordinarily observed on first exposure to new perceptual-motor tasks are attributable to differences in previous experience, the preponderance of data fits in rather well with this assumption.

Discrimination Learning. A rather extensive review of the literature of the past five years reveals no basis for changing materially the characterization emerging from Stevenson's earlier review (1963). "Work in the area is heterogeneous and fragmentary and it is not possible to reach general conclusions [p. 436]." With regard to group comparisons, the collection of earlier studies reviewed by Stevenson leaves little doubt that, as could scarcely fail to be the case, groups of retarded individuals exhibit somewhat slower discrimination learning than normals of similar CA. However, many of the studies failed to reveal significant differences.

More recent studies begin to yield some clues as to conditions responsible for the mixture of similarities and differences. Of particular importance appear to be variations in task difficulty and individual differences in previous histories and adjustment to experimental conditions which bear upon attentiveness to the relevant stimuli in discrimination problems. With regard to the former, although moderate positive correlations have generally been observed between rates of discrimination learning and both MA and IQ (e.g., House & Zeaman, 1960a), neither substantial nor consistent differences have been obtained from comparisons between normal and retarded groups with MA equated. Reviewing the collection of available studies, Zeaman and House (1967) noted that most experiments which failed to yield advantages in rate of discrimination for normal children over equal MA retardates had utilized particularly easy or particularly difficult tasks. They conclude that with tasks of intermediate difficulty and with a wide range of IQs included in the sample, "at least a low positive correlation exists between IQ (with MA controlled) and performance in visual discrimination tasks [p. 198]."

In the search for the basis of this correlation, little positive evidence has been forthcoming for hypotheses attributing it to either the performance or the associative aspects of discrimination learning as these are conceived in classical learning theories. While MA may well be correlated with motivational determiners of performance, there is little support for a motivational interpretation of differences in discrimination learning, since variation in incentives has been notably ineffective. Neither is there support for any substantial role of a general inhibition deficit. A careful examination of this possibility yielded little evidence that nonreward of responses to the negative stimulus of a discrimination problem is relatively less important as a determiner of performance for retardates than for normals. On the basis of their analyses of backward learning curves, and some experiments conducted under conditions promoting attention to the relevant dimensions of the discrimination problems, Zeaman and House (1967) have arrived at the hypothesis that even retardates, once they attend to the relevant stimulus dimension, learn visual discriminations very rapidly, perhaps often in a single trial. Some support for this idea comes from a study by Denny (1964, p. 112) who found that differences in discrimination learning between retardates and normals of similar MA disappeared when special attention was given to the handling of the subjects by individual experimenters with a view to assuring attention on the part of the subjects to the relevant cues of the discrimination problem. Zeaman, House, and Orlando (1958) found that discrimination learning by retardates was considerably facilitated by previous training in associating verbal labels with the relevant cues (colors).

Comparisons of different IQ groups with respect to discrimination reversal have yielded generally inconsistent results and only the barest start has been made toward elucidating the basis for the differences. House and Zeaman (1962) found little difference between retardates and normals with respect to the advantage of a simple reversal over a nonreversal shift, the former proceeding more rapidly for subjects of all IQ levels. A followup study by Sanders, Ross, and Heal (1965) replicated most of the conditions of the House and Zeaman study but found only a slight advantage for reversal over nonreversal shift for normals and essentially none for retardates of similar MA. However, the two studies differed considerably with respect to the amount of overtraining on the original discrimination. A later study by E. S. Ohlick (reported by Ross, 1966), utilizing only retarded subjects, demonstrated a large advantage for reversal over nonreversal shift when overtraining had been given on the original discrimination but no advantage when there had been no overtraining.

Tests for transposition of stimulus relationships to new problems following discrimination learning have, somewhat surprisingly, yielded no consistent or substantial differences between normal and retarded groups (Rudel, 1959; Stevenson & Iscoe, 1955).

Learning Set. The term *learning set* was introduced by Harlow (1949) to denote the progressive improvement in speed of discrimination learning observed over a series of problems for both animal and human subjects. The conditions for demonstrating formation of a learning set differ from those for concept learning in that the problems in the learning set series are formally unrelated. Generally the tasks are of the same general character, like discriminations between junk objects, but the positive (and similarly the negative) stimuli of successive problems share no common property which could mediate transfer of training in the usual sense.

The concept of learning set would seem to be of special importance for the understanding of mental development and mental retardation, since it arises from one of the few systematic lines of research into the way in which rates of learning are modified by previous experience. A number of early comparative studies, reviewed by Denny (1964) and Stevenson (1963), revealed a generally positive relationship between rate of development of learning sets and MA. Further, there was some indication that retardates gain more slowly than normals of equal MA, at least early in a series of problems (see also Kaufman & Peterson, 1965). Exploring the limits of this phenomenon, House and Zeaman (1958) found that retardates with IQs of about 30 made virtually no progress toward learning-set formation when tested with problems very similar to those on which monkeys had shown substantial improvement. However, with a slightly simplified situation, these investigators obtained evidence of progressive improvement even in these very low IQ subjects (House & Zeaman, 1959). Kaufman and Peterson (1958) found that retardates with IQs of 50-75, although somewhat inferior to normals, made substantial progress over a sequence of 95 object-discrimination problems. More detailed analysis showed the difference in performance between normal and retarded groups to be associated primarily with a greater incidence of perseverative errors on the part of the retardates. A number of other studies (reviewed by Denny, 1964), utilizing groups of retardates at various levels, have found substantial progress in learning-set formation for retardates, with the rate of progress being directly related to IQ when CA was held constant.

The fact that in nearly all cases retarded subjects make substantial progress in the development of learning sets even over relatively short series of problems would seem to be of far greater import than their generally lower rates of progress as compared to equal-CA normals. The formation of learning sets on discrimination problems is simply a special case of the far more ubiquitous phenomenon of "learning to learn" which has been demonstrated in a number of other situations for normal subjects. It is not obvious why it should be a matter of great priority to continue demonstrating normal-versus-retardate differences in rates of learning to learn. But it would seem to be of the greatest moment to analyze more deeply the factors responsbile for improvements in learning rates and to determine by much more prolonged and sustained investigations in a variety of situations the extent to which the relatively lower learning rates generally characteristic of the retarded may be raised by various types of experience.

Verbal Learning. Differences between normal and retarded individuals in performance on verbal tasks seem scarcely to need documentation in view of the considerable overlap in materials and tasks between intelligence tests and experimental verbal learning situations. As with normals, the most extensive studies of simple verbal learning in retardates have been conducted in paired-associate situations. Over a variety of materials, some advantages have usually been observed for normals over retardates of equal CA, the difference tending to increase over successive lists (Lipman, 1963). When the paired-associate items are not explicitly verbal in character, for example pairs of pictures, and especially when there is no time pressure, retardates with IQs as low as 50 exhibit acquisition rates almost equal to those of normals of similar CA (Denny, 1964; Eisman, 1958). In general, with verbal materials, substantial correlations are observed between IQ and errors to learn within the population of retarded individuals (holding CA constant).

When normals are matched with retarded individuals on CA, the latter show a decrement in acquisition rate, the amount of decrement increasing with difficulty of the task. However, these differences tend to disappear when MA is equated, suggesting an important role of previous learning experience.

In serial verbal learning, even higher correlations between MA and performance have been observed but, again, little difference remains between normals and retarded groups when MA is equated. For groups of equal CA, a difference in pattern of errors emerges, normals exhibiting relatively higher frequencies of errors in the middle positions of a serial list, and, conversely, a greater advantage in the end positions (Lipman, 1963). This last finding fits in with evidence from quite different tasks (which will be reviewed in a later section) tending to suggest that retarded individuals are especially deficient in the use of rehearsal and similar strategies for the facilitation of recall.

Although there have been a fair number of relevant studies, it is very difficult to determine whether there is any relationship between IQ level and long-term retention of material that has once been learned. In reviews of the literature, Denny (1964) and Belmont (1967) both arrived at essentially negative conclusions, but emphasized the difficulties of interpreting extant studies because of the very great difficulty of controlling simultaneously for amount of training and level of performance prior to the retention interval. Although further analysis with improved experimental paradigms would be desirable, it seems fairly clear that moderately retarded individuals exhibit no conspicuous long-term retention deficits as compared to normals of similar CA.

More Complex Verbal Processes. Although it might be expected on both common-sense and theoretical grounds that major differences between normal and retarded individuals should be observed in tasks involving more complex verbal and cognitive processes, documentation will evidently have to await the development of more adequate methods than are now available for analyzing these processes within either population. Studies of verbal mediation of concepts generally find deficits in retarded subjects as compared to normals of similar CA, but also indicate that the retarded can learn to use verbal mediators with specific training (Barnett, Ellis, & Pryer, 1959; Griffith, Spitz, & Lipman, 1959; Hamilton, 1966; Milgram, 1967; Stephens, 1966). On the basis of his studies of learning ability in children of different IQ levels, Jensen (1963) has concluded that,

> The habit of making verbal responses, either overtly or covertly, to events in the environment seems to be one of the major ingredients of the kind of intelligence that shows itself in school achievement and in performance on intelligence tests. Without this habit, even a child with a perfectly normal nervous system in terms of fundamental learning ability, will appear to be retarded, and indeed *is* retarded so long as he does not use verbal mediators in learning [p. 138].

Wallace and Underwood (1964) found that fewer implicit associative responses to verbal stimuli were given by mentally retarded subjects than by normals. In the normals such responses tend to interfere with paired-associate learning but aid free recall. In the retarded subjects the frequency of these associative responses seems too low to exert either effect to an appreciable extent. In general, it is quite consistently observed that mentally retarded individuals are especially deficient at retaining or transferring the use of verbal mediators even when specifically trained in a particular situation.

In the area of problem solving there seem to be almost no interpretable data available. One of the few relevant experimental studies, that of Ellis and Sloan (1959), yielded a direct relationship between rate of solving form-oddity problems and MA and also a positive relation between rate of solution and IQ when CA was held constant.

Turnure and Zigler (1964) found that mentally retarded subjects were more likely than normals to use other people's behavior as a guide to problem-solving efforts. Both retarded and normal subjects were found to become more imitative after induced failure experiences. In a later study by Achenbach and Zigler (1968) retarded subjects learned a relative-size discrimination as fast as normals under standard conditions. But if an obvious but inefficient cue were added (a light which appeared over the correct stimulus after a delay) the retarded individuals continued to use this cue much longer than normals. Achenbach and Zigler referred to this tendency as a cue-learning strategy and hypothesized that retarded individuals tend to depend on such strategies in problem situations rather than searching for intrinsic relationships. They found evidence that this tendency could be overcome by specific training. On the whole, scarcely a beginning has been made toward separating the roles of differential previous experience and possible differences in capacity variables in this area.

Summary. From the rather large volume of work dealing with comparisons between normal and retarded groups in a variety of learning situations, only a few very tentative conclusions can be drawn. It does seem clear that none of the learning or retention processes that have been analyzed in normal human learners in the laboratory differ qualitatively in the mentally defective, except perhaps at the most profoundly retarded level (below IQ of 30). Quantitative differences in rates of acquisition of knowledge or skill are generally demonstrable as a function of IQ, with relationships generally slight for simpler laboratory tests and larger for more complex ones, but tending to disappear when MA is equated.

Despite broad-scale and intensive investigation, little empirical support has been generated for the idea that specific deficits in elementary learning capacities might be found to characterize types or levels of mental retardation. However, a rather similar conception is apparently proving more fruitful in relation to relatively complex learning and cognitive processes. Individual differences in speed of discrimination learning and verbal recall appear to be importantly related, not to associative abilities but to habits of selective attention and spontaneous utilization of verbal mediators.

There are numerous indications that differences between normal and retarded groups with respect to previous learning experiences may have important bearing upon the differences in performance observed in laboratory learning experiments. Major efforts need to be directed toward determining the relative contributions of individual differences in capacity and in previous opportunities to learn over the full range from simple conditioning to conceptual learning and problem solving.

VI. *Modifying Rates of Learning and Forgetting*

The fruits of studies of mental development, including retardation and acceleration, should not simply be a collection of descriptions of individual differences, but should also provide insight into ways of producing desired changes in rates of learning or in retention. One who surveys the empirical literature on learning in relation to mental development (e.g., Denny, 1964) or, for that matter, the standard literature on laboratory studies of learning (e.g., Hall, 1966; McGeoch & Irion, 1952) cannot but be impressed at the vast quantity of work that has gone into demonstrating effects or differences which achieve statistical significance but are too small in magnitude to be of much practical import. Consequently, in this section I shall attempt to abstract from the literature bearing on each of the principal experimental paradigms reviewed above, and on selected educational settings, the studies demonstrating variables or procedures which produce large effects upon learning rates.

In this context, it is difficult to define "large" in any precise way, but, roughly speaking, I shall limit attention to variables whose effects are so substantial and reproducible that variations in them might be expected to produce readily detectable effects upon the learning rates of virtually all individual subjects. It will be convenient to follow much the same outline as in our previous review of empirical studies, in each case pointing out the implications of my general theoretical position and summarizing relevant extant results.

1. classical and avoidance conditioning

In the theoretical view developed above we assume that associations form whenever the organism is effectively stimulated by any two stimulus patterns which occur in succession and that the process is relatively independent of other circumstances except for variables which may facilitate or disturb the consolidation process. Thus variation in observed rates of conditioning must be attributed largely to degrees of optimality of conditions for bringing the behavior used as an index of conditioning under control of the CS.

Since response evocation is assumed always to involve summation of inputs from stimulus and drive sources, intensity of the US must always be a variable of major importance. In some theories, for example that of Guthrie (1952), intensity of the US is important only in so far as it is related to magnitude of the UR. In the present formulation, the principal factor is the direct relationship between US intensity and the activity of associated drive-amplifier sources. Numerous studies in a variety of conditioning situations confirm the expectation that large differences in rate of acquisition can be produced by variation in US intensity (see, e.g., Gormezano & Moore, 1962; Passey, 1948; Pavlov, 1927). We should predict that the same relation would appear even if the UR were prevented from occurring, as for example in the case of conditioning under curare, but there seem to be no relevant data available.

Over the range of values used in most experiments, intensity of the CS should be a less critical parameter. At very high intensities, of course, CSs, like any other stimuli, will activate positive or negative drive sources, resulting in facilitation or retardation of conditioning according as the systems activated by the CSs and USs are mutually facilitatory or inhibitory. At lower levels, CS intensity is relevant primarily as it is related to the probability that the learner will perceive the relationship between onset of the CS and onset of the US. In highly controlled experimental situations, where the subject can scarcely fail to be stimulated by the CS on every trial so long as it is above the threshold, over a considerable range of values CS intensity should bear little relation to rate of conditioning. However in less controlled situations, where the organism is engaged in varying activities and may not attend to stimuli of low intensity, the intensity of the CS may become a factor of considerable importance. Both of these expectations are well documented in the literature (with regard to the former see, e.g., Grant & Schneider, 1948, 1949; and Kimble, 1961, pp. 118-122; with respect to the latter, Kamin, 1965).

It will be recalled from our earlier theoretical analysis that in order for a stimulus to acquire control over behavior it is necessary that onset of the stimulus lead to some change in the organism's anticipation of reward or punishment (more technically, positive or negative drive input). If, for example, some signal has preceded a painful stimulus, then other things equal we should expect upon later occurrences of the signal that the organism will display fear, as for example by suppressing any ongoing positively motivated behavior. For this result to obtain, however, it is necessary not only that the organism experience the two stimuli in succession at some time but also that the painful stimulus occur more often in the presence of the signal than in its absence. Although it seems almost unaccountable in retrospect, this extremely important relationship was largely overlooked in the conditioning literature until the appearance of a recent series of studies by Rescorla (1966, 1968). These important experiments demonstrated clearly that for a given number of pairings of a CS and US, say a signal and a shock, probability of the occurrence of the CR upon subsequent tests could be varied over an extremely wide range by varying the frequency with which the shock occurred during intervals of time in the situation when the CS was not present. In a special case when the probability of a shock during an interval of time was the same whether or not the CS occurred during the interval, the CS was found to have acquired no control whatever over responding.

Various findings which have seemed otherwise unrelated and which have been accounted for by various independent, ad hoc hypotheses, may be interpreted in terms of this major principle. For example, it is well known that if a stimulus to be used as a CS has been presented a number of times alone at the beginning of an experiment, the rate of subsequent conditioning during a series of CS-US presentations is retarded (e.g., Harris, 1941). It is apparent that if the CS is paired with the US from its first appearance in the situation, the organism can very quickly learn that the probability of a US is greater in the presence of the CS than in its absence, whereas if the CS has appeared a number of times alone, a number of pairings of CS and US must necessarily be required before the organism can learn that the US now has a higher probability in the presence of the CS.

Another pertinent example may be the "blocking" of conditioning demonstrated in an extremely provocative series of experiments by Kamin (1965). In his basic experimental paradigm, Kamin trained rats to anticipate shock by presenting a stimulus, S_1, followed by shock a number of times in succession (a standard training procedure for the CER). Then he added a second stimulus S_2, presenting S_1 plus S_2 followed by shock on a number of trials. When, following these two phases of training, tests were given with S_2 alone, the animal's behavior revealed no evidence of conditioning to S_2. Kamin suggested that prior conditioning of S_1 had resulted in blocking the conditioning of the "redundant" stimulus S_2 when it was later added to the CS compound. Another way to look at the same Kamin study is to note that during the first phase the animal established a certain tendency to anticipate shock upon occurrence of S_1 but that the later addition of S_2 was correlated with no change in the shock and thus no change in the anticipated punishment. Thus, although the animal must have noticed the addition of S_2 to the CS compound, (as indeed was shown to be the case in some of Kamin's later experiments), the necessary conditions were not satisfied for establishment of behavioral control by S_2.

A variable which might be expected on theoretical grounds to influence the course of conditioning to a major extent, but which has been rarely studied explicitly, is the degree of reproducibility or variability of the stimulating situation from trial to trial. Other things equal, one would predict from stimulus-sampling theory (Estes, 1955; Neimark & Estes, 1967) that conditioning would be more rapid, the less the stimulus variability over trials, but that, conversely, retention of the CR once established and transfer to new situations in which the CS appears in different contexts would be directly related to variability during training. A study by Voeks (1954) utilized rather extreme measures to stabilize the stimulating situation at the time of CS onset from trial to trial in an eyelid-conditioning experiment with human subjects and obtained extremely rapid (almost one-trial) conditioning in contrast to the much slower rates observed regularly by other experimenters under usual experimental conditions. There appear to be no data available concerning retention or transfer as a function of stimulus variability.

One of the most intensively studied relationships in classical conditioning is the time interval between onsets of the CS and the US. A theoretical analysis of the effects of this variable involves a number of factors which may interact in complex ways. Firstly, other things equal, probability of formation of conditioned associations should be maximal when US follows CS onset immediately and should decrease as the interval between the two events increases, the steepness of the function depending upon the likelihood, under particular experimental circumstances, that other behavioral events which would interfere with conditioning might occur during the interval. The experimental literature (summarized by Hall, 1966, pp. 196–213) thoroughly documents the importance of the CS-US interval, since, under some circumstances, probability of conditioning varies virtually from 0–100% as the interval is varied over a range of only a few hundred milliseconds. However, the exceedingly heterogeneous results of studies in different situations must seem bewildering upon superficial examination, since the optimum interval for conditioning is very narrowly limited in some situations, as for example eyelid conditioning, whereas in others rates of conditioning are relatively independent of CS-US interval over a range of minutes or even hours.

In order to make sense of the variegated results, we should note first of all that conditioning has been observed at very long intervals (in some cases an hour or more) only when both CS and US are of such a character that stimuli having appreciable communality with either would be unlikely to occur during the interval. Adequately documented cases are those in which both CS and US are gustatory or in which the CS is gustatory and the US is exposure to x rays, inducing gastrointestinal disturbances and nausea, (Garcia & Koelling, 1966). In the situation where the index of conditioning is not a narrowly defined response but rather a persisting postural adjustment or change in level of ongoing activity [for instance, CER, passive avoidance conditioning, or "conditioned anticipation" of food (Estes, 1948)], the function relating rate of conditioning to the CS-US interval is relatively shallow, and effective conditioning can be obtained with intervals several minutes in duration. When the response has a clear preparatory function, that is, a positive or negative effect upon the efficacy of the US (as in the case of conditioned salivation with food or acid as the US), effective conditioning is obtained with intervals of the order 30 sec to 1 min (Pavlov, 1927). When the CR is a sharply defined phasic reaction, like the eyeblink, and particularly when the occurrence of this response will affect reception of the US only if it occurs during some brief interval of time, the function relating rate of conditioning to CS-US interval is steep, with the optimum value falling in a narrowly defined interval, usually of the order of .25-.50 sec.

Relations between the CS and US other than the interval between them are also important for the conditioning process. It has frequently been observed, and would be expected on theoretical grounds, that conditioning is facilitated if the CS initially evokes a response similar to, or at least not antagonistic to, the UR. Some recent studies have shown further that relationships apparently involving simply the modalities of the CS and US may be of major consequence. For example, Garcia, McGowen, Ervin, and Koelling (1968) showed that when rats were given flavored food pellets followed by administrations of x rays which produced illness, avoidance behavior on subsequent tasks was controlled by the flavor of the food pellets but not by the size. Contrariwise, when the pellets were given preceding administration of electric shock to the skin, subsequent tests revealed conditioned avoidance behavior to be a function of the size of the pellets but not of the flavor. These effects were very large in magnitude, the less effective stimulus in each case yielding essentially no evidence of conditioning whereas the more effective CS yielded substantial evidence of conditioning in a very few trials.

In avoidance conditioning, to an even greater extent than in the classical case, the interval between CS and US and the intensity of the US interact in complex ways with other conditions, variation in each of these parameters producing large effects on rate of learning in some situations and almost none in others. The theoretical principle which enables one to make sense of these findings is that for avoidance conditioning to be effective, conditions must be such that the organism can produce some clearly perceptible change in the stimulus situation by means of its behavior, with contingencies which permit it to learn that the stimulus complexes existing before and after the change are followed by the painful US with high and low probabilities, respectively.

Avoidance conditioning is difficult in some situations because, though the onset of the CS readily becomes a signal to the animal that shock is likely to follow, occurrence of the prescribed avoidance response produces no clearly perceptible change in the situation. A study by Bower, Starr, and Lazarovitz (1965) has shown that conditioning is greatly facilitated if conditions are arranged so that the occurrence of the avoidance response not only prevents the shock but also produces a clearly perceptible new stimulus which is correlated with absence of shock (and thus becomes a "safety signal" for the organism). Similarly, large differences in rate of acquisition are regularly observed between one-way and two-way shuttle situations (see, e.g., Olton & Isaacson, 1968); in the former situation the animal starts from the same end of the shuttle box on every trial and thus can readily learn that one end is associated with high probability of shock, whereas in the two-way shuttle box no such relationship exists and the animal is required on each trial to leave the end box which was safe on the previous trial and approach the one which was previously dangerous.

It is well known that, in particular in human subjects, observed rates of conditioning can be influenced to a major extent by factors entirely extrinsic to the associative process, notably those having to do with verbally induced attitudes or sets. Aspects of the experimental situation, most especially specific instructions, which lead the subject voluntarily to facilitate or inhibit the CR may produce increases or decreases in CR frequency by as much as a factor of two or three as compared to levels produced by training under more neutral conditions (Hilgard, 1938; Kimble, 1961; Miller, 1939). Similarly, manipulations of the experimental situation and instructions which lead the subject to anticipate discontinuation of the US may produce apparent extinction of the CR prior to the first nonreinforced trial (Cook & Harris, 1937).

2. operant conditioning

In view of the well-documented efficacy of operant conditioning methods for establishing control of behavior, it might seem that the operant literature should be rich with examples of procedural manipulations which produce major differences in rates of conditioning. However, it has been a tradition in this area for the acquisition process to be taken more or less for granted. Investigators learn in the laboratory the suitable procedures for producing effective conditioning in a given situation, and once having brought behavior under control by suitable reinforcement contingencies proceed to study the effects of schedules of reinforcement, drugs, or other conditions. It is difficult to find explicit controlled comparisions of the effects of different procedures upon acquisition rates.

On the basis of general experience, it appears that optimal conditions for acquisition are essentially those which one would expect on the basis of theoretical analysis. First of all, it is always essential to prepare for conditioning by insuring that the organism accepts the reward which is to be used, and that some clear signal of the administration of reward is recognized by the learner. Conditioning of so simple a response as bar pressing in the rat or key pecking in the pigeon can be tedious if the animal has not first been trained to accept food readily in the situation or if the dispenser does not produce some clearly perceptible change in stimulus conditions whenever food has been made available.

Even with human subjects, attention needs to be given to the appropriateness of the reward for the particular type of learner and situation. For example, Locke (1968a, b) found this variable to be of critical importance in obtaining verbal conditioning with moderately retarded individuals. Although previous attempts, with "good" used as a reinforcer for verbal responses, had been unsuccessful, substantial effects were obtained with tokens (exchangeable for money or trinkets) as reinforcers.

When operant conditioning is conducted in social situations with reinforcements of one individual being provided by the verbal responses of another, personality variables are bound to be of considerable importance in that the same verbal utterance may have quite different degrees of effectiveness as a reinforcer for a given individual depending upon the characteristics and behavior of the person doing the reinforcing (Binder, McConnell, & Sjoholm, 1957).

Complex behaviors must be built up progressively, the experimenter or trainer insuring at each step that contingencies are such that the organism will be effectively stimulated in close temporal succession by the discriminative cues for reward, the stimulus consequences of responding, and the stimulus changes signifying the onset of reward. With human subjects, even the mentally deficient, verbal instructions and demonstrations can be effectively utilized to further this program (Ramsey, 1963).

Some major respects in which rates of conditioning of verbal behavior are modified by different types of contingencies have been discussed in a previous section (p. 22). However, for most psychologists, the most impressive evidence of the efficacy of operant methods consists, not in these manipulations of complex behaviors, but in the establishment of simple learned behaviors in situations where other methods have conspicuously failed, for example, in institutionalized psychotic patients (Barrett & Lindsley, 1962; Lindsley, 1956), in autistic children (Lovaas, Freitas, Guilani, Nelson, & Whalen, 1967; Metz, 1966), and in the severely mentally retarded (Fuller, 1949).

3. discrimination learning

According to the stimulus–response association theory of an earlier period, one might expect that maximization of rates of discrimination learning could be achieved primarily by optimizing conditions which insure that the organism makes the correct response in the presence of the positive stimulus during training; thus the parameters demanding major consideration would be level of motivation and amount of reward. It is perhaps not entirely coincidental that the experimental disconfirmation of such expectations has been more-or-less paralleled by major changes in discrimination theory. By and large, both rates and limits of discrimination learning have turned out to be relatively independent of either motivation or amount of reward, provided only that the motivating and reinforcing conditions are adequate to maintain the behavior involved in training.

The precise manner in which the organism is brought to the point of making the necessary responses in the training situation may be of considerable import. Frequently, particularly in studies with animals, young children, or the mentally retarded, pretraining consists in first rewarding the learner for responding to both positive and negative stimuli; only after the learner responds readily to all stimuli that will be involved in the task are differential contingencies arranged so that only the correct response to each stimulus is rewarded. A result of this kind of pretraining is that mastery of a discrimination must entail unlearning of habits of responding which were rewarded during pretraining but which were classified as errors after discriminative contingencies were instituted.

An alternative to the tedious experimental extinction of nonadaptive responses is the temporary breakup of old habits, and consequent instigation of response variability, by an abrupt change in conditions. For example, Zeaman, House, and Orlando (1958) reported that after several retarded children had failed to progress on a color-form discrimination for over 200 trials, the replacement of the original positive or the original negative stimulus with a novel one led to sudden learning of the new problem.

By contrast, under a procedure first explored by Skinner (1938) and later developed systematically by Terrace (1963a, b), the animal is trained in such a way that it is rewarded for making the correct response to the positive stimulus from the outset and never develops strong incorrect responses which have to be unlearned. For example, in training a pigeon to peck a red key and withhold response from a green key, Terrace rewarded the bird for pecking at the red key until the behavior was well established, then during intervals between trials on the red key he gradually faded in the green color in such a way that the animal, which initially had no tendency to peck a darkened key, did not begin responding to the key during unrewarded intervals. By suitably programming the training, it was possible to introduce the negative stimulus, in the example the key illuminated with green as opposed to red light, sufficiently gradually so that the animal did not begin responding to it at any step. Ultimately the animal was observed to peck the red key and ignore the green one even though it had never experienced nonreward for responding to the latter.

Conspicuous features of the establishment of discriminations by training without errors are, firstly, that the total number of responses that need be made during training is generally much less than under conventional procedures, and, secondly, that at no point in training does the learner give evidence of the frustration or other emotional behaviors which frequently result from unrewarded responding.

The phenomenon of errorless discrimination learning may be closely related to the results on easy-to-hard transfer discussed in another section (pp. 161-162). Among others, Terrace (1963a, b) working with pigeons, Lawrence (1952) working with rats, and House and Zeaman (1960b) working with retardates, found that learning of a relatively difficult discrimination could be greatly facilitated if the subjects were first trained on an easier discrimination involving the same stimulus dimensions [although this qualification may not be essential (Barnett and Cantor, 1957)].

In Terrace's experiment, the animal was first trained on a very easy discrimination between the key illuminated with a particular color of light and the dark key (to which it initially has almost no tendency to respond) and then gradually transferred to more difficult discrimination between two colors of light. It is instructive to note that Terrace was unable to achieve errorless learning of a more difficult discrimination between a vertical and horizontal bar. However, this result could be achieved indirectly by first training the birds on a color discrimination, say red versus green, without errors, then superposing the vertical and horizontal bars on the positive and negative stimuli, and finally gradually fading out the colored lights.

The guiding principle in all of these studies appears to be that optimum conditions for discrimination learning require that the learner attend to relevant stimulus dimension and that the correct response be brought under control of changes to and from the positive value of this dimension at the earliest possible stage of training.

Although rates of discrimination learning seem generally to be relatively independent, over considerable ranges of values, of most of the major parameters of simple conditioning, for example, drive level, magnitude of reward, and intensity of the US, they are strikingly susceptible to variation in conditions having to do with attention to discriminative stimuli and to the relations between these and changes in the rewarding or punishing outcomes of responding.

Early studies by Muenzinger and Fletcher (1937), among others, showed that the speed of discrimination learning in the rat could be substantially increased over usual levels by such measures as the introduction of a glass barrior or a gap in the pathway of the discrimination apparatus just prior to the choice point, which forced the animal to pause and thus increased the likelihood of its perceiving the relevant stimuli at the choice point before making a choice response. Following a quite different approach, Lawrence (1949, 1950) showed that once animals had learned the relation between a stimulus property and reward while acquiring one discrimination, they exhibited an accelerated learning rate, as compared to controls, on a new discrimination involving the same positive cue, even though the response involved was quite different. Bower and Grusec (1964) showed that much the same result could be produced if the animals were simply given prior training in which the positive cue was paired with reward presentations and the negative cue with absence of reward; a similar phenomenon has been demonstrated with children by Witryol, Lowden, and Fagan (1967).

Conversely, if the individual for some reason has learned not to attend to the relevant cues of the discrimination task, rate of discrimination learning is substantially impaired. For example, Levine (1962, 1963) showed that, with adult human subjects, even a very few trials of partial reinforcement training with respect to cues which were clearly discriminable by the subjects severely crippled the acquisition of a discrimination between these cues when subsequently they were differentially related to reward and nonreward.

Comparing the learning of visual and tactual discriminations by 3–4-year-old children, Blank and Altman (1968) found both original acquisition and reversal to be more rapid for the tactual problems. Analysis indicated that this superficially surprising result is attributable to procedures which more effectively constrained the children to attend to the more subtle relevant cues in the tactual stimuli. In a second experiment, reversal of the visual discrimination was substantially facilitated by an imposed delay on each trial during which the children viewed both of the stimuli presented before making a choice, and was further speeded if the children were required to draw the stimuli they chose. Similarly, Caron (1968) found that acquisition and transfer of a visual-form concept (angularity) in 3-year-old children were substantially facilitated by procedures which led the children to respond actively to, and therefore presumably attend to, critical features of the stimulus figures.

In the case of mature human subjects, verbal instructions are the most common means of directing attention toward the relevant cues in a discrimination problem. Under appropriate conditions, the same method may prove effective even with the severely retarded. An early study suggesting this possibility (Zeaman, House, & Orlando, 1958) showed that, for children in the imbecile range, rate of learning a color discrimination was closely related to ability to name the colors. More recently, Cunningham (1965) demonstrated facilitation of visual-discrimination learning on the part of both normal nursery-school children and equal-MA retardates when specific instructions were given, calculated to direct the children's attention toward the relevant cues. This result was extended by Milgram and Noce (1968), who showed further that verbalization of an irrelevant dimension interfered with learning. Dickerson, Girardeau, and Spradlin (1964) obtained evidence that verbal pretraining on names for discriminative cues facilitated discrimination learning by retardates when names of positive cues were learned but not when names of negative cues were learned. This last result, it will be recognized, is more readily interpretable on the assumption that attentional responses are directed to cues than on the assumption that they are directed to dimensions (see pp. 170–171).

Attention to relevant stimulus properties is not a sufficient condition for efficient discrimination learning. The organism must experience in close temporal contiguity the relevant cues and the differential rewarding consequences of selecting the positive or negative values on relevant dimensions. In early studies of discrimination learning it appeared that the rat could learn a discrimination even between such conspicuously different stimuli as black and white doors only with tediously prolonged training, running to hundreds and even thousands of trials. Lashley diagnosed the source of the difficulty as being the remoteness of the discriminative cues from the rewarding consequences of choices and showed that in a suitably designed apparatus (a jumping stand in which the animal simply jumped through the door on which the stimulus pattern was displayed and received reward immediately on correct choices, or collided with a locked door on incorrect choices) the rat could learn relatively rapidly to discriminate not only between black and white but between such patterns as open and filled triangles, squares, and circles, or vertical and horizontal stripes in training series of the order of 100 trials (Lashley, 1930). Further improvements in the apparatus by Fields (1954) yielded results showing that the rat could not only learn discriminations between complex visual patterns, such as pictures cut from magazines, relatively rapidly, but could also learn and maintain a considerable number of such discriminations concurrently.

The importance of contiguity of cue, response, and rewarding or nonrewarding outcomes is not limited to the lower animals. In more recent studies Jarvik (1953) and Murphy and Miller (1955) have shown that primates are susceptible to variation in these factors, discriminations which are learned with extreme rapidity when cue and reward are in close spatial juxtaposition becoming difficult or impossible with even small separations. Similar results have been demonstrated with young children by Jeffrey and Cohen (1964).

Various means other than simple spatial or temporal contiguity can be used to aid the learner to perceive relationships between changes in relevant stimuli and changes in rewarding outcomes. For example, Bauer and Lawrence (1953) showed that discrimination learning by rats in a simple two-choice apparatus could be considerably facilitated if the colors (black or white) of the pathways corresponded to the colors of the positive and negative goal boxes. In similar vein, Shepp (1964) showed that learning of successive discriminations by retarded children could be substantially facilitated if distinctly different rewards were uniformly associated with the two positions in the modified Wisconsin general test apparatus (WGTA) and similarly that simultaneous discrimination learning could be modified over an extremely wide range according as stimulus properties of the cues and rewards were positively or negatively correlated.

Similarly striking modifications of discrimination-learning rates in the pigeon have been reported by Jenkins and Sainsbury (1969) in a recent series of studies involving the correlation of distinctive stimulus characteristics with either reward or nonreward. The situation involves two keys, one of which typically presents a small, distinctive stimulus such as a circle or square on a uniformly lighted background whereas the other key is uniformly lighted with no such feature present. When the distinctive feature is on the positive key, discrimination learning, as measured by relative frequency of choices of positive key, is extremely rapid, but when the distinctive feature is on the negative key, acquisition is severely retarded with, in some cases, no departure from chance responding evident over a long series of trials.

Detailed analyses of the particular loci on the keys at which the birds pecked revealed that rates of learning in the two cases were actually about the same, the animal learning to peck the distinctive feature when it was associated with reward and to avoid the distinctive feature when it was associated with nonreward. In the feature-positive case however, conditions were optimal for the animal to associate approach to the positive feature with increased probability of reward. In the feature-negative case, conditions were similarly favorable for the animal to associate approach to the negative feature with nonreward, but avoidance of the negative feature did not in itself increase the likelihood of selecting the positive key, since the uniformly illuminated portions of the positive and negative keys were identical.

4. paired-associate learning

On theoretical grounds, what should be the main factors in determining acquisition rates for paired associates? We shall not consider the important variables of difficulty and amount of material to be learned, since our interest in this section is with ways in which rate of learning can be modified for any specified task. Almost as important as the material to be learned is previous experience on the part of the subject which determines the extent to which the stimulus and response members of the paired-associate items function as units. In general, however, the stimulus and response aspects should not be expected to be equally critical. Whether or not it is advantageous for the stimulus members to behave as units depends upon the structure of the list or set of material being learned. Under some conditions the type of stimulus patterning required to differentiate the stimulus members of the various items in a given list might be quite different from that which would be established by previous experience with the same stimuli in other contexts. On the other hand, it is difficult to conceive of any situations in which learning would not be facilitated by interrelationships among the response members which integrate them into readily retrievable units.

Another facet of past experience which contributes to learning rate is the subject's tendency to observe or attend to differential aspects of both stimuli and responses and to rehearse stimulus–response associations during intertrial intervals. Finally, when paired associates involve verbal material, preexisting associations between aspects of the stimulus and response members become an important variable; when the subject is able to locate in memory an association which leads to recall of the response member upon presentation of the stimulus member of a pair, he can perform correctly from the outset of training with no need for forming additional associations. Both the size of the individual's repertoire of such associations and his specific skill or set toward searching for these in new situations must be important determiners of performance during paired-associate learning.

One of the most conspicuous generalizations arising from a review of the empirical literature on paired-associate learning is that the majority of the variables commonly studied, including distribution of practice, length of list, whole versus part presentation of list, and immediacy of knowledge of results, have very small effects upon rates of acquisition. By far the most important factor is the collection of variables subsumed under the notion of "meaningfulness" of the material being learned, and within this category, response meaningfulness has proved of considerably greater import than stimulus meaningfulness (Underwood & Schulz, 1960). Meaningfulness involves, of course, a relationship between the material and the individual doing the learning and consequently must depend heavily upon previous experience. Consequently observed rate of acquisition of paired associates can be taken to be a measure of learning ability only if aspects of the learners' past histories pertaining to meaningfulness of the material are adequately controlled. Conversely, learning of a particular type of material can be substantially facilitated if suitable measures are taken prior to training to increase the meaningfulness of the material for the subjects.

As in the case of discrimination learning, paired-associate learning can be modified considerably by measures which facilitate or interfere with stimulus differentiation. Sprague and Binder (1966) found that both mildly and moderately retarded children acquired paired associates much faster when the response members of items were oral than when they were push-button operations; evidently in the latter case, the motor response interfered with the subject's orientation toward the stimulus. Using a somewhat similar paired-associate task in which pictures were associated with pushbuttons, Smith, Means, and Fishkin (1968) showed that acquisition and transfer by retarded subjects (IQ 35–80) could be substantially improved if the subjects were forced to name the pictures and corresponding buttons.

For reasons that are by no means fully understood, the use of verbal labels and other verbal mediators is frequently found greatly to accelerate the acquisition of paired associates in subjects at all developmental levels. Quite possibly one function of the use of such mediators is to direct the learner's attention toward differentiating features of the items to be learned. Another may be to increase the number of associations available to support recall of a particular stimulus-response relationship.

A series of studies by Bryant (1965a, b) has shown that verbal labeling of pictures increases recall on the part of both normal and mentally retarded children. Recognition however is increased for normals only. A series of studies by Jensen and his associates has examined the role of other forms of verbal mediators in paired-associate learning by both normal and retarded adults (Jensen, 1965; Jensen & Rohwer, 1963a, b). Acquisition of paired associates by retarded adults was greatly facilitated when the subjects were instructed to make up sentences using both members of the paired-associate item. When the experimenter supplied mediators, acquisition was substantially facilitated but the advantage was lost on retesting 10 days later. Similarly Milgram (1967) found facilitation of paired-associate learning in both normal and retarded subjects by mediators, but only the normals retained the tendency to use mediators over a period of time. It is frequently observed by experimenters in this type of research that mentally retarded subjects do not spontaneously use mediators, but there seems to be little evidence as to whether they could be led to do so by appropriate training over an adequate period of time.

5. educational situations

A major portion of the current literature of educational psychology has to do with comparisons of different instructional methods in classroom situations. Since the methods compared are often generated by, or at least influenced by, theories of learning, one might expect the classroom to provide a major testing ground for the efficacy of learning theory as a guide to practice. It is my impression that this expectation is very common, and, further, that it is responsible for the negative attitude toward learning theory which develops in many educators upon observing that what he finds in this literature is almost exclusively a collection of strenuous attempts to demonstrate statistical significance of very small differences in learning rate.

I would like to suggest that the expectations are based on obsolete conceptions of learning theory and thus that the disappointments are unjustified. From the standpoint of modern learning theory, as outlined in the preceding chapters, ability to learn any particular type of task, and in particular the relatively complex ones involved in school learning, depends always to a major extent upon a combination of skills, concepts, and learning strategies, which are in turn a function of an individual's previous experience. Therefore there is no reason to expect that learning theory, as it evolves toward a higher level of scientific adequacy and rigor, should be increasingly able to dictate methods which will produce major changes in rates or limits of learning when applied across the board to groups of individuals with differing learning histories.

We should expect direct applications of learning theory to be of major effect largely as they are directed to the diagnosis and management of individual cases. Of necessity it must be difficult or impossible to assess the efficacy of learning theory at any particular stage of development by controlled experiments. It may be instructive, however, to cite a few particular examples illustrating what can now be accomplished in dealing with cases of learning disability and in suggesting some potentialities for the exploitation of learning theory in this type of application.

A second major point of contact between theory and practice lies in the development of technology. Devices such as teaching machines resulting from engineering advances do not in themselves influence rates of learning but they do permit the trainer or educator to apply more effectively whatever principles, be they good or bad, he chooses to employ.

Analyses of Individual Cases. Although an individual's failure to attain some criterion of performance is conventionally taken as evidence that he lacks some necessary capacity, appropriate analysis often uncovers some quite different source of the observed disability. In dealing with the severely mentally retarded, for example, one must be alert to the possibility that an individual's apparent inability to make a discrimination may actually reflect the experimenter or tester's failure to communicate with the subject. Some rather impressive successes in dealing with this type of problem by application of operant conditioning techniques will be discussed in Chapter XII, together with other applications of operant principles to the training of simple habits in individuals of extremely limited behavioral repertories.

It is quite likely that even at much higher educational levels cases of apparent learning or perceptual disability sometimes actually reflect failures of communication. Because of the greater complexity of the learning or testing situations, these cases will necessarily be less obvious and may often be uncovered only by careful experimental analysis.

An interesting case in point has been reported by Caldwell and Hall (1969). Tests for reading readiness in young children involve evaluation of the children's ability to distinguish similarities and differences among letters and other printed forms. In earlier work by other investigators which set the stage for Caldwell and Hall's study, reading readiness had apparently been substantially increased by training whose function was interpreted as that of leading the children to attend to distinguishing features of letters. Caldwell and Hall noted the alternative possibility that children might fail a reading-readiness test if they misunderstood or misinterpreted the concept of similarity involved (in particular, the notion that letters of similar form but different orientation should be judged to be the same). In their experimental study, Caldwell and Hall obtained substantial support for this hypothesis and, further, showed that suitable pretraining with respect to the conception of similarity involved in a test situation led to striking improvement in letter discrimination.

When standard methods fail to produce learning in an individual of apparently normal ability, the usual initial response of the trainer or teacher consists in increasing the incentives for successful performance or in instigating larger amounts of practice. However, no increase in motivation and no amount of practice can be expected to produce learning if essential constituent habits or skills are missing from the individual's repertory owing to some defect of previous experience. Several recent studies reported in the Russian literature illustrate an analytical approach which would seem to deserve wider exploitation and which should provide a particularly effective vehicle for the application of principles drawn from learning theory.

To take a particular example, Leontiev (1957) reports a study of children who had been found by music teachers to be incapable of reproducing orally their aural perceptions, sounds of given pitches, or short melodies. The hypothesis was proposed that an essential link in the normal development of this ability is the association of vocalization with pitch variations. If these associations have not been previously established, prolonged training on reproduction of melodic patterns must be expected to fail. The procedure suggested by this hypothesis was to feed individual musical notes into the subject's ear by means of earphones as he attuned his voice to the pitch of the note. Then the tone generator was turned off and the subject continued to make the sound independently. At the next step the subject reproduced the sound after an interval of up to a half-dozen seconds then proceeded to reproduce simple patterns of notes. Following this training the children were able to pass the tests which they had originally failed. Leontiev concludes,

> investigation of cases of inability to perform fully certain definite intellectual actions shows that this is only a function of an incorrectly formed functional system which can be successfully reconstructed. To do this, as experiments have shown, it is necessary to bring to light the undeveloped or incorrectly developed links, and to take people back to the stage at which the links in question are formed in their elementary external form [p. 232].

Another study illustrating the same motif involves 7-8-year-old children backward in learning arithmetic (Slavina, 1957). The particular children under investigation showed unusual "intellectual passivity," which was attributed to unfortunate earlier training. To remedy this passivity, the children were induced to enter into games involving numerical calculations. This measure seemed to lead to more positive affect and "lively cognitive activity" on the part of the children but not to appreciably greater success in learning arithmetic. Detailed analysis of the behavior of some of the individual children revealed that although they knew the names of the digits and could apparently count, they failed to grasp the conception of one number succeeding another or of the principle that any number and its successor differ by one unit. The children then were given special training in the elements of counting, manipulations with rows of sticks being coordinated with the use of numbers. With the explicit use of habits of counting thus established the children were able to learn addition and subtraction in the experimental setting. It is further reported that following this training they were able to learn multiplication in the regular classroom at an apparently normal rate.

Technologies. The most widely known application of learning principles to the technology of education is the combination of apparatus and procedures which has become familiar under the rubric "teaching machines" (Lumsdaine & Glaser, 1960). As introduced by Pressey (1926), the teaching machine was simply a device which permitted a pupil to proceed through the memorization of a set of verbal material at his own pace, being exposed to a single item of information at a time and continually testing himself by means of multiple-choice questions.

In Skinner's hands the rationale for utilizing the teaching machine was broadened by the application of operant conditioning principles and expanded into the conception of programmed instruction (1954, 1958). The advantages that Skinner saw in the teaching machine over classroom teaching were the possibilities of analyzing material to be learned into very simple, manageable stimulus-response units and arranging contingencies of reinforcement so that the pupil would have to master each successive stimulus-response relationship before proceeding to the next step in his curriculum.

One basic idea in Skinner's conception which can be of considerable import for any problem of teaching or training is that of analyzing the body of information to be acquired in a particular course in terms of the specific ways in which an individual who has mastered the material would be expected to respond differently to various questions and problem situations from an individual who has not. A second major idea, related to the conceptions of errorless discrimination learning (Terrace, 1963a, b) discussed in a previous chapter, is to break a training sequence down into sufficiently small steps so that the pupil can proceed through the sequence making very few errors and thus receiving a very high proportion of rewards for correct responding. As in other applications of operant conditioning, errors are taken to be indications of inadequacies of the program rather than defects on the part of the learner.

One important means of reducing errors is the "fading" technique, whereby the pupil is first presented with the full set of constituents of a relationship which is to be learned; then on successive occasions elements are eliminated until finally the pupil is able to reproduce the entire assemblage from memory upon presentation of some single cue (for example, to recite a poem upon being given the title or to do a problem in long division given only the divisor and the dividend). Another principal advantage of the teaching machine was conceived to be the ease of providing immediate reward, that is, arranging that each response made by the pupil in the course of proceeding through the material would be followed immediately by reward if correct and immediately by correction if in error.

The question of principal interest in the present context is whether the use of teaching machines and the associated rationale have produced any unusually successful results in bringing the performance of mentally retarded individuals up to normal levels. Though there have been a number of applications, most of these have been rather exploratory in character and it does not yet seem possible to evaluate the potentialities of the approach.

In some of the earlier reports, Ellson, Harris, and Barber (1968) and Birnbrauer, Bijou, Wolf, and Kidder (1965) report some success in teaching vocabulary to retarded children who had presumably made little progress with usual methods. Price (1963) found no advantage for teaching machines in the instruction of retarded children in addition and subtraction. On the other hand, Blackman and Smith (1964) report that their programmed instructional sequence produced substantial increases in word recognition in institutionalized retardates. By the end of their sequence the subjects were able to read phrases and sentences and exhibited relatively high retention of these skills over a 3-month interval. Similar results were reported for the acquisition of simple arithmetic skills. Malpass, Hardy, Gilmore, and Williams (1964), also working with institutionalized retarded children, report that their programmed instruction sequence led to the acquisition of word recognition and simple reading and spelling skills at a rate appreciably superior to classroom teaching. Again retention was impressively high, 75–90% on tests of ability to read words, phrases, sentences, and paragraphs over a 3-month interval. Similar results were reported by Lawson and Watson (1964).

Virtually no evidence is available in the form of experimentally controlled comparisons between programmed instruction and conventional teaching methods in the training of individuals with learning disabilities. This, is not to be taken as a criticism of investigators in this area, however, for fully controlled comparisons are of necessity extremely difficult to achieve and, further, I can see no reason to think that efforts in this direction deserve high priority. Surely the first order of business must be to find techniques which produce substantial effects on performance, then by suitable manipulations and variations within the particular approach to achieve an understanding of the factors responsible. The positive results reported to date doubtless have been achieved because of some and in spite of other features of the methods used.

On general theoretical grounds I should be inclined to believe that the favorable results obtained with teaching machines may be attributable primarily to the attention to detailed analysis of any particular learning situation, and the use of errors as diagnostic information. On the other hand I have some reservations about the heavy emphasis on immediate rewards often associated with the use of teaching machines and other applications of operant conditioning methods. Motivation and reward are important, but for effective training they need be used so that multiple relationships between actions and rewards do not simply place a substantial additional burden on the learner's information-processing capabilities.

Some of the principles of management of rewards that seem to me to be indicated by modern learning theory are perhaps more effectively embodied in a quite different bit of technology, the "autotelic environment" devised by Moore (1966) for the purpose of teaching linguistic skills to very young children. With Moore's "talking typewriter" the child receives very prompt and clearly interpretable feedback from his responses and conditions are optimized for the close association of spoken sounds, for example letters of the alphabet, which the child hears, and the corresponding visual forms which he produces on the typewriter. The child is not distracted by the administration of reward following individual responses, but attention is given to motivating him to enter the experimental situation and take part in the tasks provided.

I have seen no experimental studies of this technique, but rather detailed accounts of individual cases reported by Moore (1966) suggest considerable success at teaching reading to children at a much younger age than that at which reading instruction is usually attempted. It might be noted also that Moore makes considerable use of the device of embedding the associative learning task in a larger activity, for example, learning to read or to type letters in the activity of producing a classroom newspaper.

The latter aspect of Moore's procedure could be expected to be important for its motivational value but perhaps more important for the possibilities of inducing much more extensive and more varied repetition of learned responses than can be secured solely within experimental or school learning situations.

It is well established that the availability of learned associations is strongly influenced by the frequency with which they have been activated (see, e.g., Allen, Mahler, & Estes, 1969). Even within intensive training programs, however, the amount of repetition accruing to associations established in experimental or school situations must be much smaller than that commonly attained in everyday life situations outside the laboratory. The substantial success claimed by recent Russian writers for their programs of education of the mentally retarded in their "auxiliary schools" is attributed to an important extent to the use of "the method of modified repetition," that is the instigation of repetition and rehearsal in varied contexts, and the establishment of connections between material learned in school and practical activities carried on outside (Shif, 1969).

Vastly increased potentialities for providing practice in varied contexts within an experimental situation, and more importantly for adapting instructional programs to individual cases of learning disability, are provided by the new techniques of computer-assisted instruction which are just becoming available for experimentation on a major scale. Also, with these techniques it is possible to collect and analyze data from educational learning situations on a scale adequate to open the way to rigorous tests of the implications of learning theories in practical settings. One of the first substantial reports of a computer-assisted educational program for elementary-school students (Suppes, Jerman, and Brian, 1968) has provided evidence of an important function of short-term memory in the solution of arithmetic problems by children, thus opening up the possibility of bringing to bear upon the teaching of arithmetic results of the extensive current experimental and theoretical work on short-term memory.

The same advances in technology may, within the foreseeable future, permit the emergence of a type of research which at present is almost entirely lacking and in the absence of which only the most limited progress can be expected toward theoretical interpretation of the development of learning abilities. At present nearly all experimental studies of learning involve very short time periods, tapping the abilities of the individual only at one particular stage of development. Developmental studies provide some picture of the learning and cognitive abilities which characterize individuals at different ages which thus must be presumed to develop as a result of the haphazard sequence of experiences available within the normal range of environments. But virtually no information is available to permit the comparison of the effects of different specific sequences of experiences over appreciable periods of time on the development of motives, general behavioral strategies, or learning and cognitive abilities.

Part C

Applications of Learning Theories

VII. *Hull's Theory*

The dominance of Hull's theory in American research on learning between 1940 and the late 1950s can scarcely be overestimated. One concrete, though informal, bit of documentation results from a survey of journal literature on learning conducted by members of my graduate seminar on learning theory at Indiana University in 1955.

The students were instructed to classify each article on learning in the 1954 issues of the major American journal sources (*Journal of Experimental Psychology, Journal of Comparative and Physiological Psychology, Psychological Review*, and *American Journal of Psychology*) in terms of the theory, if any, which played a significant role in the formulation of the research problem or the interpretation of the results. Of 190 articles examined, 58% were judged to be related to some learning theory to an important extent, and, of these, 46% were formulated and interpreted in terms of Hull's system (including the "neo-Hullian" theories of Spence and Miller). Even these figures do not fully indicate the degree of influence of the system, for in many of the articles judged to be primarily empirical in orientation, such concepts as drive, reinforcement, stimulus trace, and inhibition were used in much the same way as in those more closely affiliated with the formal theoretical system.

Hull's system represented a major attempt at a synthesis of the traditions which guided a great part of American research on learning prior to 1940—associationism, conditioning, and law of effect theory. As a preliminary to discussing the representation of individual differences and the approach to problems of mental development within the system, it will be useful to review briefly some of the major concepts and assumptions. For this purpose the formulation given in *Principles of Behavior* (1943) provides a convenient point of departure.

1. major concepts and postulates

One of Hull's principal theoretical contributions is a definite and workable conceptualization of the distinction between learning and performance—more specifically, of the contributions of learning and motivational conditions to the determination of behavior at any time.

The effects of prior learning are represented in the concept of habit, including inhibition which may be viewed as a sort of negative habit. Behavior at any time is taken to be a joint function of habit and drive, according to the familiar formula

$$E = H \times D - I,$$

where E denotes excitatory tendency for some given bit of behavior, H the habit involved (that is, the associative relation between stimulus and response), D the drive condition, and I inhibition. (Throughout the section some of Hull's subscript notation is suppressed for simplicity of exposition).

It is assumed that any given index of response tendency is some function of E, that is,

$$R = f(E).$$

This formula is simply a shorthand representation of a family of specific formulas each of which would express some response measure as some specific function of excitatory tendency, e.g., $P = \Psi(E - L)$, where P denotes response probability, Ψ is the normal probability function, and L the threshold of excitatory strength; $T = k/E^c$, where T denotes response latency; $n = c'E - k'$, where n denotes number of responses to extinction; and $A = c''E - k''$, where A denotes response amplitude. Habit strength H is, in turn, assumed to be a function of a number of independent variables: t, the delay of reinforcement following response occurrence; t', the interval between stimulus and response on practice trials; w, the amount of need reduction occurring following a response occurrence; and N, the number of training trials. The function relating habit strength to each of these variables is assumed to be exponential in form so that

$$H = M \exp t \exp t'' \, (1 - \exp w) \, (1 - \exp N).$$

Thus, in the acquisition of a particular stimulus-response association, say a single vocabulary item such as the relation between the printed word *hat* and the spoken word "hat," the habit to be established is that relating S_{hat} to R_{hat}. The value of H for this habit would grow exponentially as a function of practice trials; I would probably be negligible since little work is involved in verbal responses; and D

would depend upon instructions and other relevant motivating conditions, for example, degree of anxiety. If the habit associated with the correct association were the dominant one at a particular time, then the corresponding excitatory tendency would be an increasing function of D, and performance of the correct response would be positively related to motivation. However, the rate of growth of H itself would be independent of motivating conditions except insofar as these might control the performance of behaviors conducive to practice as opposed to some interfering activity.

The principal mechanism assumed to underlie the transfer of learning to new situations is stimulus generalization. This notion is represented in the theory by the assumption that, when a habit has been acquired relating a particular stimulus to a given response, the habit strength is automatically generalized to other stimuli which are similar to the given one, the amount of generalization being a function of similarity. Thus, in the example of the vocabulary item, once the association between the printed word *hat* and the oral response is learned, generalization should be expected to occur to a very high degree to stimulus displays in which the word *hat* is printed in slightly different type, or in capitals as opposed to lower-case letters, and this generalization will permit the subject to make the correct response to this printed pattern in a variety of new situations. At the same time, generalization, though in much smaller degree, might occur to the printed word *hot*, in which case it would lead to an incorrect response. Thus, in order to build up a capability to respond appropriately, without errors either of omission or commission, initial acquisition must be followed by discrimination training in which the association between *hat* and "hat" continues to be reinforced while the association between *hot* and "hat" is unreinforced. The consequence of the discrimination training would be that inhibition, I would build up for the incorrect habit, thus reducing the corresponding excitatory tendency; the values of E for the correct and incorrect habits would diverge until ultimately the response "hat" would be made only to the correct printed word and never to the incorrect one.

In the major revision of the principles, presented in Hull's *Essentials of Behavior* (1951), the two main changes that are of importance to the present exposition were the introduction of the concept of incentive motivation (K) and a revision of the relationship between excitatory tendency and the various independent variables. The concept of incentive motivation was introduced to take account of the past history of reward associated with the given habit, K being larger if rewards had been large on previous occurrences of the response member and smaller if previous rewards had been of smaller magnitude. The formal relation between K and reward magnitude is the standard growth function, K increasing from zero at a zero value of reward magnitude to unity at some indefinitely large magnitude.

The new function relating excitatory tendency to the principal constructs is

$$E = k \times H \times K \times D - I,$$

where the parameter k subsumes the effects of other variables. A modification in the fine structure of the theory is that independent variables other than number of practice trials are now assumed to influence excitatory tendency directly rather than by way of variation in habit strength.

More recently Spence (1956) concluded on the basis of various types of animal learning data that K should enter in a slightly different way, namely

$$E = k \times H(D + K).$$

A consequence of this change, if the formalism is taken literally, is that as training on a given task proceeds, performance becomes less dependent upon drive level than in the original theory. In Hull's formulation, regardless of the stage of training, if the value of D approaches zero, so also does the value of E, and the response must cease to occur. In Spence's revision, since incentive motivation is presumed to increase over a series of rewarded trials, at a late stage in training when K has reached an appreciable magnitude, a reduction in the value of D to zero does not drive E to zero and performance, though it might be reduced, does not disappear completely.

2. interpretation of individual differences in learning ability

Now, how should we expect individual differences in learning ability to be treated within Hull's system? Hull himself addressed this problem explicitly on only one occasion, and that after the initial presentation of his system in *Principles of Behavior* (1943). In a paper in 1945 Hull sketched his view of the role of individual and species differences in a theory of behavior. In brief, his conception was that individuals differ, not in the presence or absence of various factors or mechanisms, but only in the values of the parameters representing relative weights of different factors or rates of growth of the various theoretical quantities. Thus any learner, whether normal, superior, or mentally retarded, should be expected to learn in accord with all the assumptions of the system, and a representation of his learning process would include all of the concepts, E, H, D, and so forth.

A major locus of individual differences in learning ability is the set of parameters in the function relating the growth of habit to the various independent variables of Hull's system. Thus one would expect to find a family of curves relating the growth of H to number of practice trials, with the curve rising most steeply for superior learners and most slowly for the severely retarded.

Major questions that remain to be answered by appropriate research are whether variation in the various parameters is correlated or independent and whether variations in learning ability are reflected primarily in rate constants or in asymptote constants.

Each of these questions has major ramifications. Regarding the first, it is conceivable that mental retardation might turn out to be associated with a deficit reflected in a very small value of some one parameter, as for example the rate parameter in the equation relating H to N. On the other hand, it might turn out that the parameters vary independently from one individual to another and that the retarded learner is one who happens to have an unduly large number of low parameter values.

The answers to these questions are especially important with regard to implications for training and other remedial procedures. If it were found that the mentally retarded differed from normals primarily in rate constants, the implication would be that deficits of learning ability could be compensated by sufficiently great increases in amount of training or practice. If deficits proved to be associated with asymptote parameters, such as M in the function for growth of habit strength, the implication would be that the limits of learning are lower for the retarded than for normals and that the most one can hope for is to bring the retarded individual to his maximum performance as rapidly as possible. In either event, it might be found that, within an individual, all parameter values tend to be correlated, so that if rate of growth of one habit is unusually low so also will be the rate of growth of others, or it might be the case that parameters relating to different types of habits (as, for example, verbal versus nonverbal) are unrelated. It should be especially emphasized that the theory does not purport to deal directly with these questions but only provides a framework for research which might yield answers to them.

Another major possibility which can be meaningfully expressed within Hull's system is that the mentally retarded differ from normals, not with regard to rates or limits of habit acquisition, but with respect to other factors which determine performance. To the extent that this state of affairs obtains, considerable analysis may be needed to determine just how variation between individuals in particular factors would affect skilled performance. To take one example, the slopes of stimulus generalization gradients might well vary among individuals. A shallower generalization gradient would imply greater transfer of learning from a training task to a slightly different test situation, and this transfer might be thought to be advantageous in most educational situations. However, as illustrated in the example of the vocabulary item discussed above, generalization can also lead to errors. Thus, while it might be more adaptive for the organism to have a shallow generalization gradient with respect to the acquisition of fears, it would probably be more efficient to have steep generalization gradients in situations involving discrimination learning.

With regard to inhibition, one's first thought might be that since inhibition tends to counteract habit strength in the determination of excitatory tendencies, a low level of inhibition would tend to accompany superior performance. However, in many situations the acquisition of a skill involves the overcoming of errors, which in Hull's system would be represented by the building up of inhibition associated with undesired habits. Thus there would seem to be no sound reason for expecting an overall difference in level of the inhibitory variable between learners of different ability levels.

Within Hull's system, in contrast to classical associationism, there are clear reasons to consider the possibility that large differences between individuals in apparent learning ability might actually result from performance factors unrelated to learning in the usual sense. The most obvious candidate for this role would be nonspecific drive level D, which combines multiplicably with habit strength to determine performance. No amount of training can be expected to lead to major changes in performance on any given type of task if the D level is so low that excitatory tendencies do not exceed threshold. But even more important than the nonspecific effect of D on activity level is its presumed function in discrimination and selective learning.

According to Hull's conception, any situation involving problem solving or discrimination learning necessarily also involves alternative responses on the part of the individual and thus a number of competing habits, some of which are correct and some incorrect from the standpoint of the requirements of a particular task. Within the theory, if some habit strength exists for the correct response and some habit strength for one or more incorrect responses to a given situation, the probability of the correct response depends not only on these habit strengths but also on the drive level. Since the difference in habit strengths would be multiplied by *D*, whether or not a given difference in habit strengths will yield a suprathreshold difference in excitatory tendency depends on the current drive level.

However, the role of *D* in the acquisition of skills is double edged, for if at some stage in training habit strength of some incorrect response is greater than that for the correct response, amplification of the difference in habit strengths by a high *D* level will simply increase the probability of an error. Thus, individual differences in *D* should be expected to lead, not to an across-the-board difference in skilled performance, but rather to a pattern of differences changing with stage of training.

Incentive motivation *K* might be an even more important variable than *D* for the investigator concerned with individual differences in mental development. Since the value of *K*, unlike that of *D*, is assumed to grow as a function of rewarded trials in a learning situation, individual differences in level of *K* would be less likely to yield a mixture of positive and negative effects on the acquisition of skills. At the beginning of training on any new problem, when habit strength for errors might be greater than those for correct responses, the level of *K* would presumably be low, and so also would be its effect as an amplifier of these differences in habit strength. At a later stage in training, when habit strength for the correct response would normally be greater than that of any particular incorrect response, an increased level of *K* would have only beneficial effects since, as a multiplier of the difference of habit strengths, larger values of *K* would yield larger probabilities of correct responding.

Thus individual differences which resulted in some individuals having lower rates of growth or lower asymptotes of incentive motivation could have widespread effects on the acquisition of skilled performance. In fact, since different habits must be assumed to have different growth rates within any individual regardless of his overall level, it seems quite likely that individual differences in the growth rates or asymptotes of incentive motivation might be a much more likely source of broad deficits (or the converse) in apparent learning ability than individual differences in rates or limits of habit growth.

Another reason why the variable *K* might prove an especially attractive focus of research lies in the possibilites of overcoming performance deficits by remedial or therapeutic measures. To the extent that mental retardation is a matter of individual differences in parameters of habit growth, it might well be that little could be done to overcome deficits by modifying training procedures. However, there would be far more reason to think that much could be done to overcome or compensate for differences in levels of incentive motivation. One might expect that investigators working within the framework of Hull's theory would direct major efforts toward unraveling the extent to which individual differences in intellectual functioning are attributable to habit, drive, and incentive variables.

Now, to what extent have the potentialities of Hull's theory begun to be exploited in the directing of research on mental development? In order to evaluate progress in this respect, it will be useful to consider separately the relatively direct applications of the Hull-Spence theory and theoretical formulations of a general or special character which represent outgrowths or extensions of Hull's system.

3. direct applications of the Hull-Spence theory to research on learning ability

The initial attempts to apply Hull's concepts to the analysis of mental retardation have been discussed in a systematic review by Cantor (1963), to which readers may also refer for a summary of the Hull-Spence system somewhat more comprehensive than the one given above.

With respect to the most direct and straightforward application of Hull's interpretation of individual differences, namely the determination of differences between normal and retarded subjects in parameters of the habit-growth curve, Cantor's review discloses but a single example up to 1963. The study in question was one by Ellis, Pryer, and Barnett (1960), who reanalyzed data from some previously published experiments on serial verbal learning by groups of mentally retarded individuals (averaging about 17 years of age with a mean IQ of about 64) and normals of about the same age (with mean IQs of approximately 103). With Hull's concept of reinforcement interpreted in this situation as occurrence of the correct response followed by knowledge of results, probability of correct response was plotted versus number of previous reinforcements for a pooled group of 118 defectives and a group of 75 normals. The curves were well fitted by Hull's habit-growth equation with a suitable choice of parameter values, and it was found that the curve for the normals required both a higher rate parameter and a higher asymptote parameter than that for the defectives. Thus the result is compatible with the conclusion that the normal subjects differed from the retardates in both the rates and limits of habit acquisition. However, interpretation of the result is complicated by factors which perhaps help explain why there have been few similar studies.

As indicated by our previous analysis, one cannot actually determine the constants of the habit-growth curve from empirical learning curves of this sort, since the response measure, in this case, probability of correct responding, is a function of excitatory tendency which is in turn a function of drive and incentive factors as well as habit strength. The extent to which, for example, the observed difference between the two mean empirical curves could be a function of differences in drive level cannot be ascertained from these data.

In order to get further insight into the theoretical locus of differences between the performance curves of the normal and retarded learners, it would be necessary to do studies of this sort, not in isolation, but as part of a series of experiments in which different variables were manipulated. It should be possible, for example, by suitable manipulation of the incentive conditions to determine the extent to which differences between the performance curves should be attributed to drive or incentive-motivational differences. Further relevant information could be obtained by using a paired-associate task and, after the learning functions have become asymptotic, interchanging stimulus-response assignments among different items. At the outset of the second phase, the dominant habit with respect to each item would be one involving an incorrect response and any difference in drive level in favor of the normals would be expected to increase their probability of errors rather than to increase their probability of correct responding. The system has substantial resources for experimental manipulations of this sort which might be expected to be of considerable diagnostic value but which have not yet been exploited in relation to research on retardation.

Most of the applications of the Hull-Spence theory in the area of retardation have attempted, not actually to determine the parameters of theoretical functions, but rather to make comparisons between groups of normal and retarded subjects on single response measures which might reflect differences in theoretically significant variables, provided all of the relevant factors have been held constant in a comparison. With regard to stimulus generalization, Cantor's (1963) review includes but a single study. Barnett (1958) adapted his experimental situation from one used originally by Brown, Bilodeau, and Baron (1951). The subjects were required to learn to respond to the middle light in a row of 11 by operating a response button. Following a series of training trials, the subjects were given test trials involving single presentations of all the lights in the display. Generalization gradients were obtained, with the tendency to respond to test lights decreasing as a function of distance from the training light in the fashion qualitatively implied by Hull's theory (or any other) but with the forms of these gradients very similar for the normal and retarded groups. Little can be concluded from this one set of data, of course, especially since the stimulus dimension involved was position and it is well known that speed of learning simple positional discriminations differs little between normal and moderately retarded individuals.

Within the framework of Hull's theory, there is perhaps no special motivation for pursuing this matter intensively. However, it might be noted that in a different context, Zeaman and House (1963) have postulated that a major difference between retarded and normal learners may be localized in probabilities of attending to different stimulus dimensions. If their idea is correct, it might be expected that generalization functions on some stimulus dimensions would differ in slope as a function of learning ability. If, for example, retarded children on the average had lower probabilities of attending to color than normals, they would presumably be less likely to exhibit a generalization decrement when a test stimulus differed in color from a training stimulus and thus would exhibit shallower generalization gradients on that dimension, even though they might not differ from normals on generalization for position.

Applications of the concept of inhibition are complicated by a number of factors, not the least of which is a sharp difference in definition of this concept between Hull's original formulation (1943) and the later major reformulation of the theory by Spence (1956) which has influenced many applications of the system in various areas, including mental retardation. According to Hull, inhibition is a direct consequence of responding, the amount of inhibition accruing per response being a function of the amount of work involved in the given effector activity. Inhibition is assumed to be built up by a constant increment upon each occurrence of the given response and then to die away exponentially to zero as a function of time between responses. Thus there is room for individual differences both in the amount of inhibition generated per response and in the rate of decay.

It is not entirely obvious just how differences in learning ability should be related to these parameters. On the one hand, a higher buildup of inhibition per response would lead to a decrement in responding even during continued reinforcement and thus it would be deleterious to performance in many kinds of routine tasks, for example, the operation of a station on an assembly line. However, this factor might be interpreted as a performance variable unrelated to intelligence or learning as these attributes are usually conceived.

In many situations, particularly those which require the learner to find a correct response by a process of trial and error, a greater increment in inhibition per response might be adaptive since it would lead to more rapid weakening of initially strong but incorrect responses and thus facilitate the emergence of correct behavior. If an increased rate of buildup of inhibition were compensated by an increased decay rate, the individual having this combination would not be penalized on repetitive tasks except under conditions requiring high rates of repetition of the same response. Some studies comparing performance of normal and retarded individuals on simple tasks provide evidence of just such a pattern. For example, Siegel and Foshee (1960) compared groups of normals and defectives on a task which allowed the subject to extinguish the signal light each time it appeared by pressing any one of four switches. The defective group showed less response variability from trial to trial; this result, taken at face value, might suggest a smaller buildup of inhibition per response for the subjects of lower intelligence. A complication of this study however was the much higher average chronological age of the defectives.

In a study reported by Ellis (1970) defective and normal individuals of about the same chronological age were compared on a pursuit-rotor task involving blocks of massed trials separated by a rest period and additional training either 1 or 28 days following initial training. The greater gains in performance over the longer rest periods for the normal subjects led to the conclusion that they must have built up greater amounts of inhibition per response during the massed-trial sessions.

A similar result was reported by Jones and Ellis (cited in Cantor,

1963) although the latter study failed to confirm the associated prediction that the normal and defective subjects should have been differentially affected by the degree of distribution of practice during the work sessions. Possibly, however, the prediction that the two groups should have differed in performance during training sessions as a function of distribution of practice is not well founded since, as we have indicated above, normal and defective subjects might well differ on both the rates of increment and the rates of decay of inhibition. There seem to be no results in the literature which would enable one to tease out these aspects separately.

In Spence's formulation (1956) the effects of work are not given explicit attention and the notion of inhibition is embodied in two concepts, the inhibition resulting from nonreward (and presumably related to the commonsense notion of frustration) and the inhibition of delay observed in certain Pavlovian experiments. Cantor (1963) feels that applications of these concepts to problems of mental development might be more fruitful than the use of Hull's conceptualization, but notes that almost no relevant work in that direction had been accomplished.

Very little work seems to have been directed toward the problem of working out the way in which parameters having to do with the various determiners of habit growth vary as a function of intelligence. A study by Jacobs (1950) compared institutionalized normals of different ages and institutionalized defectives in a task involving different delays of rewards. Though immediate reward was beneficial to performance in all groups, the absence of interactions indicated that the effect of delay was similar in all groups.

Studies of incentive motivation have been rather exploratory in character and lead primarily to the impression that little is yet understood concerning individual differences in preferences for different incentives and the ways in which these can be manipulated. A study of the effects of different reward magnitudes in multiple-choice discrimination learning by defectives reported by Blank (1958) provides no information concerning the relation of this variable to individual differences in learning ability, but does confirm the expectation from Hull's theory that beneficial effects of increased incentive motivation should be manifest primarily at the points in training when correct responses become dominant (that is at the higher habit strengths).

Research directed toward evaluating the role of differences in drive level in learning of individuals of different ability has to date been primarily limited to assessment of anxiety level as measured by the Taylor manifest-anxiety scale. A major and largely unsolved problem in this line of research is the difficulty that the responses to the Taylor scale must necessarily involve reading comprehension to an important degree and thus can scarcely be independent of intelligence level. In fact, some research has indicated that scores on this type of test may be as importantly related to IQ as to emotional or drive level as these are usually conceived. Following a review of the literature in this area Cantor (1963) suggested that it is unlikely that defectives as a group differ greatly from normals with regard to anxiety but at the same time proposed that more effort needs be directed toward determining whether high levels of anxiety might be an important source of interference with learning in the mentally defective.

A notable deficiency in Hull's system is the absence of any mechanism for handling phenomena of retention in comparable detail to those of acquisition and the associated absence of any conceptual basis for differentiating short- and long-term memory.

The potential usefulness of Hull's system for research on mental ability is sharply limited by the fact that this theory is based primarily on animal learning data and has been interpreted in only an extremely narrow range of situations involving verbal and conceptual behaviors of human beings. Consequently, it is to be expected that other theories will have to be called upon to guide research of the types generally considered to be of most practical importance in connection with mental development, especially that having to do with the acquisition and use of language. Perhaps the greatest potentiality of the theory lies in its function as a conceptual tool for analyzing the roles of learning processes and motivational variables in the determination of performance.

VIII. *Ellis's Stimulus-Trace Theory*

Perhaps the best example of a limited theory of mental deficiency elaborated on the basis of a single concept drawn from Hull's theory is Ellis's stimulus-trace hypothesis (1963). Ellis's idea appears to be that a major difference between normal and mentally defective individuals lies in a short-term memory mechanism. The hypothesis may in part have arisen by exclusion since mental defectives seem notably inferior to normals in the kinds of performance ordinarily thought of as having to do with memory, yet formal experiments rarely exhibit appreciable deficits in long-term retention for material learned in the laboratory. Other relevant considerations include the common use of digit-span tests of short-term memory in intelligence test batteries and similar use of delayed-reaction tests as a means of detecting "intellectual" deficits in brain-injured animals.

Ellis (1963) pointed out that the trace concept appears in nearly all types of behavioral theories, though in varying guises. In gestalt theory (e.g., Köhler, 1929) a memory trace is basic for retention as well as for such phenomena as time errors in psychophysics. Hebb's neurophysiologically oriented theory of behavior (1949) assumes two trace mechanisms, one of which is a brief reverberatory trace persisting after stimulation and presumed to provide the basis for immediate memory. Drawing primarily upon Pavlov (1927), Hull (1943) developed the notion of stimulus trace in some quantitative detail and utilized it as a basis for the interpretation of trace conditioning, serial-position effects, temporal conditioning, and the like.

Ellis's version of the trace concept was not spelled out in detail in his original theoretical paper (1963) but appeared to be primarily a variant of Hull's stimulus-trace postulate. In a later revision (Ellis, 1970) the connection with Hull's theory seems largely to have been lost and the concept has been redefined as a short-term memory mechanism. In view of this change in orientation, it will be convenient for our purposes to review the 1963 presentation first and then the later revision and extension.

1. background and assumptions of the 1963 formulation

The essentials of the original concept were that, following the occurrence of any stimulus, a trace s_t, persists in the nervous system for a period of the order of a few seconds during which it can provide the basis for a delayed response and can become conditioned to the responses evoked by other stimuli. The magnitude and rate of the decay of the trace were assumed by Ellis to be a function of "central nervous system integrity." Thus central nervous system defects of a sort to produce defects in learning ability were assumed to be associated with reductions in the magnitude of the trace from the normal level and also with an increase in the rate of decay, so that the duration of control of behavior by any given stimulus would be less for the subnormal than for the normal individual.

Ellis took a number of types of experimental evidence as support for the idea that individual differences in the stimulus trace are important determiners of performance on learning tasks. One of these had to do with the serial-position curve in verbal learning. Although some early research (e.g., McCrary & Hunter, 1953) suggested that the form of the serial-position curve, with the maximum occurrence of errors near the middle of a list, was unrelated to ability levels, some later studies [including those of Noble and Fuchs (1959) and Barnett, Ellis, and Pryer (1960)] indicated that faster learners had a relatively greater advantage on items near the beginning or end of the list and made disproportionately more errors on the middle items.

This last finding conforms nicely to expectation based on Hull's interpretation of the form of the serial-position curve. Stimulus traces should, on the one hand, permit responses from early items to perseverate and cause errors on later items; but also these traces, diminishing in magnitude as a function of time following occurrence of an item, should cause the effects of reward at the end of the list to exert graded effects on items in relation to their proximity to the end. Therefore, of two otherwise similar individuals, the one with the stronger or more persistent stimulus trace would be predicted to conform more closely to the pattern found for the superior learners in these latter studies.

One difficulty with this source of support is that other more recent studies (e.g., Girardeau & Ellis, 1964) have not found the predicted differences. Another difficulty is that factors other than the stimulus trace may well be more important determiners of the form of the serial-position curve. In particular, the primary effect, that is, the reduction in the level of errors toward the beginning of the list, may be related to an important extent to rehearsal tendencies. Thus, individual differences in rehearsal strategies, which would be expected to lead to differences in performance on many types of learning tasks, could well lead to the more sharply bowed serial-position curve sometimes found for faster learners.

Similar comments apply to evidence from delayed-reaction studies. Over a very long period investigators in a very different context have felt that the delayed-reaction experiments should be especially diagnostic in relation to intellectual functioning. There is, to be sure, some gross correlation between performance on a delayed-reaction task and phylogenetic level, and in some instances brain-damaged animals show decrements in delayed-reaction performance. However, little success has attended attempts to localize the source of delayed-reaction decrement in damage to areas of the brain thought on other grounds to be important for associative processes. On the other hand there is some reason to think that defects in delayed-reaction performance by brain-injured animals have something to do with a deficiency in the ability to inhibit responses to relevant stimuli during the delay interval—the animals giving an impression of the same "distractability" which observers frequently find to characterize the performance of mentally retarded children.

Experimental studies comparing delayed-reaction performance of normal and mentally defective children have yielded mixed results, some studies indicating a correlation between MA and performance and some yielding no significant relationship. As in the case of serial-position phenomena, this rather inconsistent pattern of results might be more easily interpreted on the assumption that the differences in performance level are associated with differences in habits or strategies of rehearsal and other activities which can serve to bridge the delay interval rather than in some simple property of the nervous system as implied by the stimulus-trace interpretation.

Perhaps the most direct bit of evidence providing the setting for Ellis's original hypothesis was the finding by Berkson, Hermelin, and O'Connor (1961) that the duration of blocking of the alpha rhythm following presentation of a visual stimulus was significantly greater in normal than in mentally defective individuals. There is no logical connection between electroencephalographic data and the concept of stimulus trace as defined in Hull's system, but one must concede that the finding of differential durations of alpha blocking is suggestive of a relationship between intellectual level and some type of trace activity which persists following stimulation.

2. tests and implications of the trace theory

In any event, the nature of the evidence which suggested the possible importance of a stimulus-trace process in the analysis of learning deficits is not of great import. As Ellis indicated (1963, p. 145), the primary test of this concept should have to do with its fruitfulness in yielding predictions regarding such matters as the relative difficulty of different types of learning problems between normal and defective individuals. To this end Ellis and his associates have conducted a number of studies which test implications of the trace hypothesis. One of these was a study by Ellis and Zimmerman (cited in Ellis, 1963, p. 147) in which a modified delayed-response situation was utilized. Presentation of a colored light in the single upper window of a stimulus panel initiated a trial. Following a delay interval a row of lights of different colors appeared in windows at the bottom of the panel. The subject's task was to select the light from this set which matched in color the sample that had appeared at the beginning of the trial. As anticipated, performance declined as a reasonably orderly function of delay interval in a group of mentally defective adult subjects, and there was a modest relationship between MA and performance at the longer delay intervals.

However, it should be noted that over 10 days of training there was improvement in performance on the part of all of the subjects, and rather drastic improvement in some cases, so that by the tenth day some subjects' performance was approaching constancy over the range of delay intervals used. This finding does not particularly support the idea that the initially steep decrements in performance over delay interval were attributable to individual differences in a stimulus-trace parameter representing some relatively stable aspect of the nervous system. On the other hand, they fit quite well with the suggestion mentioned above that individual differences in delayed performance depend to a major extent on corresponding differences in activities such as verbal rehearsal.

A rather more direct test of the trace hypothesis was undertaken in a study by Terrell (cited in Ellis, 1963, p. 148) involving a simple reaction-time task with differing intervals between a warning signal and the stimulus to which the subject was instructed to respond. For half the subjects the warning signal remained on until the reaction-time stimulus occurred, whereas for the other half it went off after 1 sec, leaving an interval which would have to be bridged by some trace process on the part of the subject. As expected, the reaction-time function for a group of normal subjects declined only slightly as a function of the duration of the warning interval and was virtually unaffected by the on-off versus persisting warning signal conditions. Reaction time for a mentally defective group was higher at all values of the warning interval than that for normals and, more importantly, was much higher in the on-off than in the persisting warning signal condition.

The latter gross qualitative result confirms Ellis's expectation on the basis of the trace hypothesis. However, these data should perhaps be interpreted with some caution in view of the unusual finding that reaction time declined as a function of the interval between the warning signal and the reaction-time stimulus over the range 2-4 sec, in contrast to the classical result of an optimal interval in the neighborhood of 2 sec.

A rather direct implication of the stimulus-trace hypothesis is that in well-controlled learning situations, such as paired-associate learning, the interval between stimuli should be important in determining relative rates of learning of normals and mental defectives. Ellis cites two studies which appear to fit in quite well with his interpretation. Vergason (1964) compared groups of normal and retarded children, 12-16 years of age, on a simple paired-associate learning task under conditions which eliminated overlearning by dropping items from the list after they had been responded to correctly. Rates of original learning and relearning after 30 days were very similar for the two groups. Presumably there was no basis for differences in stimulus traces to have operated. However, in a study by Blue (1963), groups of normal and retarded subjects were trained in a paired-associate task under conditions in which the stimulus and response members of the paired-associate item appeared simultaneously and under conditions in which there was a delay of 2.75 sec, between the stimulus and response members. The normal subjects were virtually unaffected by the delay, but the retarded subjects learned much more slowly under the delay condition. This finding would appear to deserve further investigation since it has generally been found that paired-associate learning is relatively independent of intratrial time intervals over a wide range of values.

3. reformulation in terms of memory processess

On the whole, tests of the original formulation were encouraging for the idea that a specific short-term memory deficit might characterize the mentally defective, but not for an identification of the deficit specifically with the stimulus-trace concept as developed in Hull's system. During the same period in which the work discussed above was conducted, a major acceleration of investigative activity concerning short-term memory appeared in the work of Broadbent (1958) and Peterson (1963), among others. The similarities in methods and results of these investigations and those of some of the studies designed to test the trace hypothesis became apparent and suggested the possible fruitfulness of continuing the latter line of research within the theoretical framework developed by Waugh and Norman (1965) and Atkinson and Shiffrin (1968).

In general outline, this family of theories analyzes the learning of verbal materials in terms of operations of information storage and retrieval without attempting to reduce these to stimulus-response terms. Thus this theoretical framework takes a form much like the flow diagram of a computer program. The first stage involved in learning any particular set of material is conceived to be an attentional process, the term being used in much the same sense as in the theory of Zeaman and House discussed below. Stimulus items attended to are assumed to be automatically registered in an immediate-memory system from which information is rapidly lost as a function of time, all traces of the input being lost within a few seconds unless some mental activity on the part of the learner transfers the material to a long-term memory-storage system. Ellis (1970) proposes distinguishing three phases in the storage process rather than the two (short- and long-term storage) assumed by Waugh and Norman, and by Atkinson and Shiffrin. In Ellis's schema the first phase of storage is termed primary memory and has the properties indicated above for short-term storage in the other theories. Primary memory material may be transferred to secondary memory and then to tertiary memory.

Secondary memory is assumed also to be a transient phase, but with a slower decay rate than the primary memory system; presumably it is responsible for retention over intervals of a few seconds, as for example those involved in digit-span tests. Tertiary memory corresponds to the notion of long-term memory storage in the other theories. Transitions from primary to secondary and from secondary to tertiary memory are assumed to occur as a result of rehearsal activity on the part of the learner.

The original stimulus-trace notion would seem to be translatable into the hypothesis that a primary source of memory deficit in retardates should be found in primary memory capacity. However, Ellis departs from the spirit of the original theory at this point and hypothesizes rather that there is essentially no difference between normal and mentally defective individuals in primary-memory capacity. Further, on the basis of the slim supply of available evidence on long-term retention for verbal materials in the retarded, he assumes no difference between normal and retarded individuals with respect to tertiary-memory capacity. The principal loci at which major differences between normals and defectives are expected to be found lie in the secondary-memory system and in the rehearsal activities leading to transitions between systems.

4. research deriving from the revised theory

In order to obtain systematic data relevant to this analysis which would be comparable for normal and mentally defective individuals of various age levels, Ellis developed a serial-exposure apparatus similar in many respects to an experimental situation used in a series of studies by Atkinson and his associates (Phillips, Shiffrin, & Atkinson, 1967). In Ellis's situation the subject sits in front of a console which presents a row of nine circular windows. By means of a set of projectors behind the display panel, numerals can be exposed in any of the windows individually in any desired order and for any prescribed exposure durations. In the standard experimental paradigm, a sequence of digits, usually a random permutation, is chosen in advance by the experimenter, and these digits are exposed one after another in the row of windows from left to right. For a given experiment, there is a constant exposure time for each digit and a constant interval between exposures of successive digits. Thus, for example, the digit 7 might appear in the first window for 1 sec and then disappear; following a 2-sec interval, the digit 3 might appear in the second window for 1 sec; then following a 2-sec interval, the digit 9 in the third window; and so on. A retention interval follows the offset of a display in the last window and then a test digit appears in a single window centered at the top of the panel. The subject's task is to indicate in which window of the row of nine the test digit had appeared by touching the proper window, thus activating an electrical contact leading to the recording system.

Presumably an individual's ability to locate the position where the test or "probe" digit appeared should reflect short-term memory relatively independently of recall processes and should permit the independent investigation of exposure durations, time intervals between exposures, and time between the final exposure and the test. Preliminary results reported by Ellis and Hope (1968) suggest that these objectives are quite well met. Their data on college students indicated that rate of presentation mainly affects memory for the early and middle positions in the series but not for the terminal positions, and that exposure duration and interval between exposures are approximately interchangeable in their effects.

The obvious interpretation in terms of the theoretical schema would be that increasing the presentation rate reduces opportunities for rehearsal and thus reduces the transitions from primary to secondary memory that are needed for effective retention of the earlier positions in an exposure series. A delay between the last digit of a series and the presentation of the test stimulus primarily affects accuracy on the last exposure position, as would be expected if the perfect retention of the final stimulus normally observed at short test intervals is principally attributable to the primary-memory system.

Comparisons between performances of normal and retarded individuals in this situation have yielded a striking pattern of results, and one of special interest in that very similar performance of normals and defectives and very large differences are both observed in the same situation (Ellis, 1970). In the first of a series of studies, college-student subjects were compared with a group of retardates of about the same age whose mean IQ was in the neighborhood of 60. Sequences of trials were run with 0- and with 2-sec intervals between stimulus presentations and with an immediate test at the end of each exposure series. The test data revealed large differences between the normal and retarded subjects at both exposure rates. However, the pattern of the differences is of more interest than the magnitude. When serial-position curves, that is the proportion of correct responses at each of the nine positions in the display, were plotted, very large differences appeared between normals and retardates at the early positions but no difference at the last position. These results seem rather negative for the earlier stimulus-trace notion, since memory for the final position should yield the data most closely reflecting a simple trace process.

In terms of the revised conceptual schema, one must conclude that normals and defectives are strikingly similar in primary memory. The principal question is whether the large differences observed at the earlier positions can be localized in the secondary-memory phase of the general model. Since the normal subjects proved to be considerably more influenced by presentation rate than defective subjects, and since the differences appear largely at the earlier positions in the sequential display, Ellis (1970) has proposed a "rehearsal strategy deficiency hypothesis" as a successor to his earlier stimulus-trace hypothesis. The essentials of the new hypothesis are that memory capacity, at least in the primary and tertiary phases envisaged in his schema, does not differ appreciably between normal and retarded individuals; but the retardates are characterized by a deficiency in the rehearsal strategies which are required to transfer information from primary to secondary and secondary to tertiary memory storage.

Some aspects of the new hypothesis have received considerable support from additional experiments in the serial-exposure situation (Ellis, 1970). When a group of retarded individuals was run for 12 days in the same basic situation but with .5-, 1-, and 2-sec exposure rates for different blocks of trials, no appreciable effect of the rate variable appeared. There was a learning effect over the 12 days of experience, and it is perhaps significant that the improvement appeared only in scores on the earlier positions of the serial display, which would be compatible with the notion that the progressive improvement represented some modest success in improving rehearsal strategies.

When exposure rate was held constant and delays of 0, 5, 10, or 15 sec were interposed between the final stimulus of the exposure series and the test stimulus, the recency effect (that is accuracy at the last stimulus position) was found to decrease considerably more as a function of delay for retarded than for normal subjects. For neither group was there any systematic trend in accuracy at the earlier positions as a function of delay, tending to confirm the hypothesis that performance at the earlier positions reflects mainly a longer-term memory storage system which can be quite sharply distinguished from the very transitory storage which is reflected most clearly in performance at the final position.

The general conclusions from this series of studies would appear to be, firstly, that the retarded individuals have a substantial deficit in the ability to retain information from stimulus displays over intervals of the order of 0-20 sec. Further, when the display is presented serially, the principal deficit for the retardates appears in the earlier positions. Variation in presentation rate produces large effects at the earlier positions for normals but little or no effect for retardates, suggesting that the normal individuals take advantage of slower presentation rates to engage in rehearsal activities which improve retention of information from the earlier presentations of the sequence. These quantitative results, together with observations of retardate performance reported by Ellis, indicate that the retarded subjects engage in very little rehearsal activity.

5. comment

Interpretation of these findings in terms of the theoretical schema raises numerous problems which call for further analysis. Firstly, consider the data regarding retention of information presented at the last serial position. It is clear that upon immediate test normals and retarded subjects, within the range tested to date, yield equal performance at virtually 100% retention. If a test is delayed, the retention curve falls much more sharply for the retarded subjects. One possible interpretation, and the one evidently favored by Ellis, is that the first point on the retention curve for the final position represents primary memory and the remaining positions secondary memory, so that one might conclude that normals and retardates do not differ with respect to the primary-memory process. However, the data available at present do not clearly select this interpretation from a number of alternative possibilities.

Perhaps the most obvious alternative hypothesis is that the short-term retention function is an inverted S curve, relatively horizontal for a short interval following the stimulus exposure, then falling more steeply for a time before the transition to the final negatively accelerated portion. This possibility could be checked out by further experimentation in which shorter test intervals were used. However, it must be noted that no experiments of this type can localize the differences between individuals of different ability in a memory storage or decay process representing purely retention of information from the final stimulus of the series. For in all cases this presentation has been preceded immediately by the others in the series and thus there is a clear possibility of the type of interference customarily associated with the notion of proactive inhibition. That is, the retention function for the final position may be influenced by the information input at earlier positions. If so, the divergence between the final-position retention curves for normal and retarded individuals might represent the emergence of proactive effects. In order to get at this possibility, it will evidently be necessary to make comparisons of retention functions for the final position with the number and spacing of earlier presentations in the sequence as an independent variable, and perhaps to investigate conditions in which the stimuli at earlier positions differ qualitatively from that at the final position.

One particularly interesting possibility is that the differences between normals and retardates in the terminal-position retention curve might be attributable entirely to the fact that the normals have stored information from the earlier positions more effectively, so that it interferes less with retention of material from the middle positions. In this event the differences as a function of ability level in this situation might be entirely reducible to differences in ability to transfer information from short- to long-term memory storage rather than to differences in storage capacities at any level.

The idea that differences in short-term retention as a function of ability level may depend to a major extent upon differences in rehearsal activity is attractive both because it is relatively open to investigation, as distinguished from hypotheses concerning differences in memory capacity, and because it entails possibilities of mitigating deficiencies in performance if ways can be found to improve rehearsal strategies on the part of defective individuals. However, before a start can be expected in the latter direction, numerous questions need to be answered by further research.

Apparently retarded individuals rehearse less in standard learning situations. Is the same true of nonlaboratory situations such as those involved in the acquisition of language or in school learning? In either event, why does the retarded individual engage in less rehearsal? Is he unable to rehearse effectively? If so, is this a matter primarily of an innate limitation on his ability for such learning or is it a deficiency in earlier learning which might somehow be overcome?

When the retarded individual does rehearse, does the rehearsal activity produce the same effects on memory storage as the same activity in an individual with normal learning ability? This last question is especially critical with respect to practical implications, for there would be no reason to engage in strenuous searches for methods of teaching rehearsal strategies to the mentally defective unless there were reason to believe that increased rehearsal on their part would improve intellectual performance.

IX. *Denny's Elicitation Theory*

Although the approach which Denny (1966) terms *elicitation theory* is, perhaps, not a direct outgrowth of Hull's system, it fits into the same general framework and shares both the advantages and limitations to a considerable degree. The conceptualization of behavior is strictly that of stimulus and response, with learning occurring by stimulus substitution, as in the theories of Pavlov and Hull. But with respect to the basic condition for learning, Denny follows Guthrie (1952) in assuming that simple contiguity of stimulus and response is both necessary and sufficient.

The principal function of reward in a trial-and-error situation, as, for example, maze learning, is that it evokes approach to the location of the goal. Then simply through association by contiguity, approach responses become conditioned to other cues near the goal and the process works backward over successive trials until exposure of the animal to any portion of the pathway will initiate a sequence of movements leading to the goal. The termination of pain is assumed to exert its effects by eliciting relaxation, which in turn reduces the likelihood of occurrence of competing responses that might otherwise become conditioned to stimuli present in the situation at the given time and interfere with or displace the response which led to termination of pain.

Experimental extinction of learned habits is assumed to proceed strictly by counter-conditioning. That is, a response which has been learned to a certain stimulus will continue to be made upon recurrences of the situation until some other response is learned in its stead.

The important concept of an elicitation hierarchy has much the same properties as the habit family hierarchy in Hull's theory. At any given time an organism is assumed to have a hierarchy of possible responses to any given situation. Over time or over repeated trials in the same situation a primary process termed "satiation" is assumed to occur, temporarily decreasing responsiveness to the repeated stimuli and thus leading to temporary changes in relative strengths of responses in the hierarchy. The shifts in the hierarchy produced by conditioning and counterconditioning are assumed to be lasting changes. Stimulus generalization from one situation to another is assumed to occur in much the same manner as in Hull's theory but the importance of stimulus sampling receives more emphasis in Denny's formulation.

On the basis of an earlier review of the literature on learning studies involving the mentally defective (Denny, 1964) and, no doubt, of his own more recent experimental work with severely retarded subjects, Denny suggests that the main defects in learning on the part of the mentally retarded are attributable to limitations on duration of attention and inability to delay response to a stimulus, deficiencies in capacity for symbolic behavior, and a broad deficit in inhibitory processes. He further suggests that the last named, the inability to inhibit responses, may in turn be responsible to a considerable extent for the other types of deficiency. In application to problems of mental retardation, Denny's theory is very similar to that of Hull. It is directly applicable primarily to the learning of simple habits. The types of procedures and the types of analyses that would be suggested by Denny's and Hull's theories would seem to differ largely in matters of emphasis, Denny giving more explicit attention to the role of competition among alternative responses in all types of learning situations.

On the basis of these assumptions, together with firsthand experience in training situations, Denny proposes the following as the major guiding principles which need exploitation in training the mentally retarded.

1. elicitation of the correct response

The first concern in every training situation must be to evoke the desired response from the learner in the presence of the stimuli which are later to signal it while at the same time minimizing the occurrence of incorrect responses. Several means for achieving these ends have been developed in various experimental contexts. Most important, perhaps, is the use of "crutch" cues and "fading" techniques (Terrace, 1963a, b). To illustrate, suppose that a child is being trained to use the words "light" and "dark" appropriately to describe stimulus cards of different shades of gray. If discrimination is difficult, one might begin by putting a black patch on the darker card; then, after this relatively easy discrimination is mastered one might gradually reduce ("fade out") the black patch over a series of trials by steps sufficiently small that the learner never begins to make errors until, ultimately, correct responses are being made to both shades of gray. Another technique suggested is the use of physical barriers to prevent the occurrence of incorrect responses during early stages of learning in trial-and-error situations.

2. immediate knowledge of results

In either learning experiments or training situations knowledge of results following correct responses ordinarily is conveyed by means of rewards which, among other things, indicate to the subject that his response was appropriate. Following incorrect responses, knowledge of results involves an indication of the error, such as saying "wrong." In terms of elicitation theory, immediate knowledge of results following errors is clearly important if the procedure allows the subject to make a correction since this would insure the elicitation of the correct response in the given stimulus situation. In a trial-and-error situation where the task involves finding a route to a goal, rewards following correct responses are clearly important in terms of elicitation theory since they evoke approach responses in the vicinity of the goal, and these responses may in turn become conditioned to more remote stimuli. When the responses are nonspatial, and especially when they are verbal, it seems less clear how to interpret the function of reward in this system. To the principle of immediate knowledge of results, Denny adds that of differential feedback. It is not entirely clear to this reviewer how the two concepts are to be differentiated except, perhaps, that differential feedback may serve simply to increase the number of stimulus elements in a situation which become associated with the desired response.

3. stimulus generalization and positive transfer

In harmony with the ideas of Guthrie and of stimulus-sampling theory (Estes, 1959), Denny emphasizes the importance of increasing the likelihood of transfer of training to new situations by giving training in a variety of contexts. For example, in teaching the name of an object, essential aspects of the training are to elicit the correct name in the presence of the object in various background situations and to use various instances of the same object (for example, a variety of tables if the word "table" is being learned). Transfer is expected to increase if this training is spread over a number of sessions rather than being highly concentrated in time, since distributing the practice increases the likelihood of changes in stimulus context and therefore of the association of more different stimuli with the correct response.

Further, in order to establish appropriate usage in verbal behavior, it is important that a particular response be repeatedly associated with a particular stimulus while other features of the situation vary randomly. If, for example, a child were taught to respond to a particular command with the training always given under precisely the same circumstances, learning would appear to be rapid, but the effects of the learning would disappear as soon as the child was tested in another situation.

4. motivation (elicitation hierarchy)

As in all learning theories, motivation is recognized to be important, but in the elicitation approach the emphasis is upon external incentives which may be expected to evoke vigorous behaviors from the learner rather than upon internal conditions of drive or need. Variety of stimulus presentation is expected to have as one of its benefits a slowing of satiation processes which might otherwise be expected to lead to reduced attention to stimuli relevant to the learning task. Rewards are considered important, not simply as indicators of correct responses within a particular segment of a learning task, but also as a means of evoking approach responses toward the learning situation as a whole.

5. sequential building

This undoubtedly important principle, which in Denny's presentation is only cursorily described, relates to the strategy of beginning with things which a given child finds it easy to learn and building by easy steps toward more difficult goals, taking advantage of positive transfer of training whenever possible. It would seem that a learning theory should have much to say of a rather specific character about the nature of sequential building, the advantages of different types of sequences, and the like, but these problems seem not yet to have been dealt with explicitly in the case of elicitation theory.

Only with respect to the first of these proposed conditions do Denny's recommendations differ substantially from those that would follow from most other classical learning theories. The use of "crutch" cues and of fading techniques can undeniably be effective, but also the use of these methods can involve considerable hazards. As will be illustrated in the following section, the use of a fading technique in a discrimination problem can result in the learner's becoming fixated on an incorrect response pattern. It is essential that the fading be done in such a way that the learner is not led to attend to an irrelevant stimulus dimension (see pp. 79-80). Further, the use of barriers, physical or otherwise, to prevent incorrect responses during early stages of learning seems intuitively appealing but can involve undesirable and unexpected side effects. For example, it is well known in maze learning experiments that an animal which has been prevented from taking a particular pathway by means of a barrier, and thus coerced into following a correct pathway to food reward over a series of trials, may have extremely high probability of taking the incorrect path the first time the barrier is removed (see, e.g., Denny, 1957; Estes & Schoeffler, 1955; Zeaman & House, 1951). A possible solution to this difficulty, which seems not to have been studied experimentally, is to eliminate a barrier gradually, by analogy with the fading technique. This possibility needs both an experimental and theoretical analysis.

Denny has applied his theoretical approach to retardate training, not be a series of experiments in which it is compared to other theories, but by developing a simple training situation and working intensively with a small number of severely retarded children over a considerable time (Denny, 1966). His teaching device is a multiple-choice apparatus which permits simultaneous presentation of several stimuli, for example, pictures, on a display panel and permits the child to respond to these by operating a response bar. Correct responses are indicated by illumination from within the knob at the subject's end of the response bar. At the beginning of training, the child is simply taught to respond to simple commands such as "go to green" by moving the bar until it touches the green stimulus patch, crutch cues being supplied by the experimenter initially in the form of gestures or manual guidance. Then appropriate responses to increasingly complex commands and simple concepts having to do with color, form, spatial relations, and the like are taught in a regimen guided by the principles given above. Although no formal quantitative assessment is given, it appears that considerable progress has been made in teaching a variety of simple behaviors and concepts to several children with IQs in the range of 20-30. Denny indicates optimism regarding the future possibilities of more expanded programs following the same principles.

X. *The Social Learning Approach*

Typically the retarded individual differs from normals not only in his cognitive functioning in laboratory experiments or in school learning situations but also in his reactions to other people, his ability to function in a social group, and in the dispositions, habits, and attitudes customarily associated with character and personality. At present it is scarcely possible even to begin analyzing the development of personality and of social interactions in terms of the same theories that have been applied to the more elementary cognitive processes. Furthermore, proposed new theories at this level do not meet the same standards of rigor and specificity of prediction. Nonetheless, it is as essential at the more complex as at the more elementary levels to guide and organize research by theoretical concepts.

Of the currently influential theoretical approaches to personality development, the one most amenable to discussion in the same context with theories of learning and memory is that which has been termed the "social-learning approach." The principal basis for our present examination is the review by Cromwell (1963).

1. social learning theory

The immediate conceptual basis of the social learning approach is Rotter's "social learning theory (1954)." This formulation is not a learning theory in the same sense as those of Hull or Spence but rather a composite of concepts drawn from the theories of Lewin (1935), Hull (1943), Miller (1959), and Tolman (1932) and applied to the interpretation of personality development and the social functioning of the individual.

The principal theoretical concepts drawn from learning theory are reinforcement value and expectation. The primary assumption is that the tendency for behavior to occur in any given situation is a function of the individual's expectation of reinforcement in that situation and of the value of the reinforcement. The notion of value is somewhat more extended than that involved in the concept of reinforcement or reward magnitude in learning theory. Reinforcement value reflects not only the value of the given outcome relative to some need or motive but also the individual's expectation that the reinforcement will lead in turn to other reinforcements and of the values of these. Expectation of reinforcement in any situation depends partly on generalization from other experiences. Further it is assumed that behaviors are organized into clusters or systems in such a way that different clusters of behavior are interchangeable if they lead to the same rewards.

Thus, the social learning approach constitutes, in effect, an outline for a theory together with some guiding ideas as to the types of concepts and relations that may prove important. I believe it would be a mistake to attempt to evaluate it in terms of its present ability to lead to specific predictions of behavior. However, its function in guiding research in the area of mental retardation is of considerable interest. It is to this aspect that Cromwell's review gives major emphasis, summarizing the preliminary application of the approach to a number of studies and sketching some of the revisions in provisionally formulated concepts that have been suggested by these applications.

2. research in "social learning"

Preliminary studies by Cromwell and his associates (Ringelheim, 1960; Starkman & Cromwell, 1958) explored ways in which the mentally retarded individual's learning is affected by previous successes and failures. Attempts to demonstrate positive or negative effects on subsequent speed of learning proved essentially negative. Also, only rather inconclusive results were forthcoming from studies intended to evaluate the notions that retardates are selectively impaired in their ability to conceptualize themselves as failing or not failing on a task or that retardates more than normals tend to reward themselves verbally rather than responding to actual rewards.

The question arising rather directly from consideration of the basic ideas of the social learning approach is that of whether the retarded individual has a lower general expectation for success and a higher generalized avoidance tendency than the otherwise comparable normal. If the answer to the first part of the question is affirmative and groups of normal and retarded subjects are matched for initial performance on a task on which correct performance is to be rewarded, then the retarded groups should improve relatively more than the normals over subsequent trials. The retarded subjects, beginning with a lower expectation of reward, would gain more from the experiences with the actual rewards obtained. Conversely, if training is given with punishment for errors, then retarded individuals should be affected less than normals since they initially have a higher expectation of failure.

It must be noted that, as in other applications of the social learning approach, the predictions are by no means unambiguous. If the levels of general expectation at the beginning of a particular experiment differ for retardates and normals because the former are slower in building up such expectations as a function of reward or punishment experiences, expectations might then be affected less than those of the normals by the rewards received in the particular experimental situation. In any event, Heber (1957) attempted to test these predictions in a simple reaction time experiment. The results were confirmatory for the analysis with respect to data from success trials, the reaction times of the retarded subjects decreasing more rapidly over trials than those of normals. However, an unexpected trend appeared for the failure trials, on which the reaction time for the normals went much below that for the retardates and then regressed somewhat late in the series. The latter finding, confirmed by Gardner (1966) in a card-sorting task, suggests that perhaps the normal individuals have learned to respond to initial failure in a task by increased effort whereas the institutionalized retardates have learned to respond to failure by avoiding the task situation so far as possible.

Other experiments attempting to bring out consistent differences between normals and retardates with respect to avoidance tendencies have yielded rather inconsistent results and led to the formulation of what Cromwell (1963) terms a developmental hypothesis. The gist of the latter is that the very young child reacts only to punishments and rewards *per se*; that is, he is influenced simply by the extent to which they affect his immediate drives and motives, lead to pleasure or pain. With age the child learns the relationships between his own behavior and consequences and comes to realize the degree to which he can control the occurrence of reward or punishment and therefore the experience of success or failure. As a consequence the older child develops motivation to engage in activities that permit him to demonstrate his ability to succeed and learns to avoid situations where his feeling of competence may be threatened.

A study by Bialer (1961) attempted to obtain confirmation of the developmental hypothesis by studying normals and retardates of various ages in two task situations. The first was a repetition-choice situation.

The subject was given two puzzles to complete, and was allowed to complete one (being informed of his success) but was interrupted in the course of the other (this being described to him as a failure). The subject was then given the choice of returning to the failed task or to the completed one. The other experimental task was a choice situation permitting the subject to choose between a smaller immediate and a larger delayed reward. Bialer found that higher CAs in both normal and retarded children and to an even greater extent higher MAs were associated with a greater tendency to return to the incomplete (thus initially failed) task and to choose the larger delayed reward over the smaller immediate reward. These findings were taken to support the notion that the individual of higher MA is less dependent upon the immediate effects of rewards and punishments *per se*. Rather, he tends to characterize these events to himself in terms of success and failure and is influenced by them in accord with his relatively generalized and persistent habits of responding to successes and failures.

Continuations of this line of research by Miller (1962) and others were principally concerned with comparisons of groups which differed in a variable termed "locus of control." This concept refers to the tendency to perceive success and failure as results of external agencies as opposed to reflections of the adequacy or inadequacy of the individual's own efforts. Relative standing on this variable was evaluated by means of a paper-and-pencil task developed by Bialer and Cromwell (1960).

For groups of normal children, Miller (1961), using a serial-learning situation, found no differences between groups tending toward external locus of control (ELC) and groups tending toward internal locus of control (ILC) under conditions of success emphasis, that is, learning with rewards for correct responses but with no punishment. However, a large advantage for the ILC group appeared under failure emphasis, that is, learning with penalties for errors but no rewards, and under neutral conditions with neither rewards nor punishments. A shift from failure to success emphasis produced much the greatest facilitative effect on the ELC groups. Thus, this study seemed to show that it is possible to characterize subjects by an individual-difference variable which has considerable predictive value with respect to the individual's dispositions to respond favorably or unfavorably to reward or punishment, success or failure. The major questions of how locus of control develops in the child and whether differences in this variable are of great import for the study of mental retardation have not yet been directly attacked.

3. a tentative formalization

The conclusions from the various studies conducted within the general framework of the social learning approach have led Cromwell to a set of postulates and theorems which he believed might be useful in extending this area of research. Examples of the postulates follow:

> I. The initial system of motivation of the human organism is to avoid noxious events and to approach satisfying events . . .
>
> II. The cumulative associations of the outcomes of goal-approach and threat-avoidance events with one's own behavior evolve and become conceptualized over a period of time into a system of motivation to demonstrate one's own behavioral effectiveness . . .
>
> III. Independent of developmental level, the magnitude of the earlier system versus the later system determines the relative sensitivity (responsivity) of the individual to cues of hedonistically positive or negative circumstances versus cues relating to success and failure (degree of behavioral effectiveness) in the individual, respectively (Cromwell, 1963, pp. 78-80).

Examples of "theorems" follow:

> I. Verbal responses expressing an internal locus of control increase as a function of age.
>
> II. The tendency to choose success (behavioral effectiveness) under hedonistically negative circumstances increases as a function of age.
>
> III. The tendency to avoid failure under positive hedonistic circumstances will increase as a function of age.
>
> IV. To the extent that mental age is a measure of the conceptual development of the individual, its magnitude will be more related to the locus of control and other success-approach tendencies than chronological age (Cromwell, 1963, pp. 79-80).

Much as in the case of Rotter's (1954) earlier presentation of "social learning theory," it does not seem sensible to this reviewer to evaluate Cromwell's collection of postulates as though it were directed primarily toward formal theory construction. The postulates are not securely grounded in empirical results, but rather represent trends that Cromwell and his associates believe to be discernible in a mass of frequently conflicting findings. The postulates contain many undefined terms and the theorems are simply suggested implications and are not logically derived from the postulates. The spirit in which Cromwell seems to have put forth the set of postulates is rather to characterize the degree of progress toward theory construction in this area and to aid in the direction of continued research.

4. comment

Although it is a sociological fact that investigators in various areas of psychology tend strongly to develop their own disciplines in relative isolation from others, one cannot help being struck by the possibility that the kinds of research most needed in further exploitation of the social learning approach might be greatly furthered if there were more utilization of progress in related disciplines.[2] For example, during the past decade, experimental and theoretical investigations of learning and motivation at both the behavioral and physiological levels have yielded major advances in understanding some of the mechanisms and processes involved in the control of behavior by reward and punishment. Surely progress toward understanding the development of social learning would be accelerated by more effective utilization of this body of relevant information and theory.

The investigation of personality development in the mentally defective, as well as in normals, must be a major long-term research goal. Nonetheless it is hard to avoid the conclusion that more methodological groundwork is needed before one can realistically expect this area to begin yielding constructive results comparable to those already demonstrable in the studies of more elementary learning and cognitive processes. None of the studies covered in Cromwell's extensive review has reported any large effects upon the performances of mentally defective individuals from the manipulation of success or failure experiences or the other variables that are taken to be important in the social learning approach.

On the positive side, the indications that personality variables which can be independently scaled may predict an individual's responsiveness to such major determiners of performance as reward and punishment suggest the possibility of important long-term benefits from continued exploitation of this approach. It seems that the major potential contribution of the approach must be in locating and studying the experiences during the earlier development of the child which lead to the establishment of motives for achievement, reward values, and responsivity to success and failure.

---

[2] Some substantial progress in this direction is reported in a theoretical essay by Gewirtz (1969) which appeared after the present volume was in press.

XI. *Field Theory*

Although numerous studies of performance and learning in the retarded have been conducted under the influence of the viewpoint associated with "field theories" in psychology, it remains difficult to elucidate just what field theory is or how it is applied to the direction of research. In general, the view of learning within gestalt theories has been that an individual's performance is determined largely by the way in which he perceives his environment, and that perception is in turn determined by organizational principles having to do with goodness of figure, perceptual closure, relations between figure and ground, and the like. Memory is essentially a residue of perceptual experiences, and memory traces undergo changes in the course of time in accord with similar principles of organization.

Early interpretations of mental development in terms of gestalt principles are largely associated with Lewin and his students. Lewin (1935) and Kounin (1941) elaborated a conception of progressive differentiation of an individual's perceptual and cognitive activity with age and interpreted mental deficiency as retardation in this differentiation process with the consequent "rigidity" of cognitive and perceptual processes. The concept of rigidity seems not to have been fully spelled out operationally, but has to do with tendencies to react to problem situations by falling back on perseverative or stereotyped behaviors.

The idea that the maladaptive behavior characteristic of the mentally deficient is a direct result of retardation in a developmental or maturation process has never found appreciable support in experimental evidence, and has been extensively criticized by Zigler (1962), who interprets the data seeming to support this conception as reflecting complex motivational factors. Zigler proposes that institutionalized retardates have higher generalized avoidance tendencies toward adults than normal individuals and at the same time higher desires for social approval, of which they are characteristically deprived. Zigler suggests that these strong sources of motivation conflict with or override those conducive to effective learning and problem solution and lead to rigid or stereotyped behaviors.

Neither approach to the loosely defined collection of symptoms of intellectual malfunctioning subsumed under the term "rigidity" has given rise to a great deal of fruitful research. The developmental conception is almost immune to investigation until developed in much more detail, and even then would have to derive support largely from the exclusion of alternative interpretations. The importance of motivational factors in the performance of the retarded is probably difficult to overestimate. With regard to the motivational interpretation, it seems all too easy to demonstrate characteristic differences in motivational patterns between people living normal lives and the mentally deficient living in institutions or in other severely limited environments but almost insuperably difficult to begin uncovering the role that motivational factors may have in the etiology of mental deficiency. As will be seen in a later section, the presently more feasible task of finding ways of modifying motivational conditions in the direction of facilitating adaptive behaviors and improving intellectual functioning has been undertaken largely within the operant conditioning approach.

In a systematic review of research on retardation that has been conducted under the influence of a field theory, Spitz (1963) finds some support for this approach in laboratory studies conducted to examine developmental trends in such processes as perception of illusions and reversible figures. The single concept from gestalt theory most prominent in this line of research is that of *cortical satiation*, originally developed by Köhler and Wallach (1944) to denote the buildup of cortical inhibition consequent upon continued stimulation. In a number of studies, Spitz and his associate have found support for their hypothesis that mentally retarded individuals are slower than normals in both the buildup and decay of cortical satiation as reflected in, for example, speed of locating hidden figures, rate of reversing reversible figures, and optimal intervals for successive masking in tachistoscopic perception (Spitz & Blackman, 1959; Spitz, 1959; Spitz, 1969).

On the basis of these studies Spitz (1963) has proposed the following four "postulates" to characterize mental retardation within the framework of field theory:

1. Changes of state of groups of cortical cells in response to stimulation are slower for retardates than for normals.
2. Following offset of stimulation the return of cortical cellular activity to the normal state is slower for retardates than for normals.

3. In the retardate it is more difficult to change once established cortical patterns into new patterns as a result of changes in stimulation.
4. In the retarded there is less spread of activity from stimulated cortical cells to the surrounding cortical field.

The immediate implications of these postulates are presumed to be: (*a*) The mentally retarded individual should learn less in a given number of trials than the normal individual; (*b*) the retarded should forget relatively less of material once learned; (*c*) the retarded should exhibit poorer transfer of learning to new situations; (*d*) the retarded should exhibit higher thresholds for such phenomena as apparent movement.

It is not difficult to see that these postulates represent a somewhat plausible translation of a number of commonly observed experimental findings into terms having to do with cortical cells and the like. However, unless much more explicit rules can be provided relating the assumed cortical phenomena to behavior, it is difficult to see that this type of formulation could be expected to serve as more than an heuristic aid in research for individuals who find it intuitively appealing.

In a later review, Spitz (1966) goes rather further toward reformulating some of the ideas drawn from gestalt psychology in terms compatible with those of other learning theories and more nearly translatable into empirically testable propositions. A new series of investigations starts from the gestalt view that memory is aided by the isolation of material from context and by organization into meaningful, cohesive units in accord with much the same principles that govern organizational aspects of perception. Within this framework one is led to inquire concerning the extent to which mental deficiency may be due to the inability of retarded individuals to organize material in memory in ways which resist decay or distortion with the passage of time.

A convenient experimental situation for preliminary analyses was provided by the standard free verbal recall experiment. Studies by Bousfield (1953), and Tulving (1962), among others, have shown that in recalling lists of objects or words, normal adults tend to recall the items in groups or clusters associated with meaningful categories, that the extent of such clustering tends to increase with practice in free recall, and that there is a positive correlation between the extent of clustering and the amount recalled. These findings led to the hypothesis (Miller, 1956; Spitz, 1966) that repetition may be important primarily in providing the individual with an opportunity to group and otherwise organize material in memory.

Early studies of free recall in relation to IQ, using pictures or familiar words as stimuli, revealed little obvious variation in the amount of recall or amount of clustering over a considerable range of IQs. Later work, mainly using verbal stimuli, established that educable retardates both recall and cluster less than normals of equal CA. When MA is equated, differences in recall disappear; however, the retarded groups tend to exhibit less clustering if the measures of clustering are corrected for "categorical intrusions." This last term refers to the tendency to make errors of commission on recall tests by giving words which were not in the lists presented but belong to the same meaningful categories as words which were in the list.

Spitz was led to a particular set of studies by the observation that college students' superior recall performance is accompanied by very high tendencies to group and to cluster items in recall as compared to normal children or retarded individuals of similar CA. On the assumption of a causal relation between clustering and recall, it seemed possible that recall in retardates might be substantially improved experimentally by measures leading them to do more grouping or clustering. Two procedures suggested by this hypothesis, one involving presenting items grouped in meaningful categories and the other guiding the subject's recall by requesting him to give items by categories, both led to very large increases in recall performance on the part of children with IQs of 60-70.

With respect to paired-associate learning, recall strategies are not as obvious, but the interaction of meaningfulness (or difficulty) of material and IQ as predictors of performance suggests that the retarded individuals are at a disadvantage in producing mnemonic devices to facilitate recall. Following out this idea, Jensen and Rohwer (1963a, b), working with adults of IQ of 50-60, produced substantial increases in rate of acquisition by instructing the subjects to form sentences relating the stimulus and response members of paired-associate items. And Jensen (1965) found that retardates given specific instructions to use this type of mnemonic device performed better than normals of equal age who had learned without the aid of special instructions.

Even memory span for digits, a subtest in many intelligence scales, has proven susceptible to the same type of manipulation. Spitz (1966) found that the recall of digits by retardates was substantially improved by instructions to group the digits (a technique evidently generally used spontaneously by normals).

On the basis of the encouraging results obtained with even small amounts of instruction in the use of mnemonic devices and organizational strategies, Spitz has arrived at the suggestion, to my mind a most promising one, that long-term studies need to be undertaken in which retarded individuals are given sustained training of this kind in order to ascertain whether substantial improvements in their capabilities for intellectual functioning might be effected.

XII. *The Operant Conditioning Approach*

Though not a theory in the technical sense, the body of doctrine associated with operant conditioning procedures and concepts has become extremely influential in guiding a major segment of research upon animal performance in relation to learning and motivation and the extension of the experimental results to the practical control of behavior. This system, originally formulated by Skinner (1938), falls in the stimulus-response-associationistic tradition and differs from that of Hull mainly in placing less emphasis upon the inference of underlying states or processes from performance and more emphasis upon the teasing out of empirical relationships between observed behaviors and controlling variables.

1. concepts and principles

The principal systematic independent variable of the system is *reinforcement*, a reinforcement being any consequence of a response which leads to an increase in probability of that response upon future recurrences of the same situation. Reinforcers are defined simply in terms of their effects upon behavior, with no attempt being made to seek independent definitions in terms of physiological mechanisms such as need reduction or the activity of brain centers. It is assumed that for any organism there exist events having the property of acting as reinforcers (else the organism would not survive). A reinforcer is expected to exert similar effects upon different responses in the same situation and similar effects upon the same response in different situations. Thus the concept of reinforcement is by no means circular. Once a reinforcing operation is identified for a given organism, the system specifies how it can be used in the future to modify behavior. The aspect of behavior altered by reinforcement is said to be *response strength*, or simply likelihood of responding, this property being most directly manifest in the rate of responding in any situation where the organism is not paced by external events.

Within this system one does not speak of the acquisition of knowledge, nor even of associations, but simply of modification in behavioral tendencies as a result of reinforcement history. Skilled performances of various sorts are assumed to develop by way of discrimination and differentiation. If a response is originally made indiscriminately to the different stimuli present in an experimental situation, as to black and white cards in a simple discrimination problem, the response is brought under the control of the stimulus property by differential reinforcement. For example, one may reinforce the organism by presenting food whenever it makes a given response in the presence of the black stimulus and withholding reinforcement when the organism makes the response in the presence of the white stimulus. If the organism originally makes the response, say operation of a response key, in a variety of ways, and one desires to train the organism to make the response always with a particular force, the procedure is to make reinforcement contingent upon responses having forces falling within a specified range and withholding reinforcement whenever responses occur with forces outside the specified range. Always a critical concept is contingency—reinforcements must be made contingent upon responses of a specified character occurring only in the presence of specified stimuli or stimulus properties.

Auxiliary principles of major import for training are *reinforcement schedule* and *shaping*. A schedule consists in a set of rules specifying the precise conditions under which an organism will receive reinforcement for a given response. A vast variety of schedules has been studied in detail for several simple operant conditioning situations, in particular bar pressing by the rat (Skinner, 1938) and key pecking by pigeons (Skinner, 1950; Ferster & Skinner, 1957). If the experimenter's or trainer's objective is to build a habit that will persist over a considerable period even when reinforcement may not be available, the training schedule is arranged so that the animal frequently has experiences of responding without immediate reinforcement and has an opportunity to learn that if it persists in responding under such circumstances further reinforcements may be forthcoming. In more systematic terms, the response under investigation becomes associated with stimuli occurring on unreinforced occasions as well as with stimuli present only on or following occurrences of reinforced responses.

Shaping behavior by successive approximations is a procedure that has evolved for training a response in a given situation even though the organism originally has no tendency to make the given response, so that under simple trial-and-error routines the response would never be rewarded and thus never strengthened. The idea is to observe the organism's behavior in the given situation and determine the class of responses which it does emit that resemble in some way the one which is to be trained. For example, if a pigeon is to be trained to peck a key, it may originally exhibit no tendency to do so but may nonetheless make movements in the direction of the key and, on other occasions, may make pecking movements. The training program then utilizes a changing series of reinforcement contingencies in which the animal is first reinforced whenever it moves, then is reinforced whenever it either makes a pecking movement or moves in the direction of the key, and finally is reinforced only when it actually pecks the key.

Because of the emphasis upon the actual modification of behavior as opposed to the inference of underlying processes, the operant conditioning approach would seem to be a natural framework for at least some types of research on the training of the mentally retarded. Indeed, substantial research programs have been under way for a number of years exploring this motif (e.g., Bensberg, 1965; Bijou, 1966; Bijou & Baer, 1967; Sidman & Stoddard, 1966).

2. approach to mental retardation

The general viewpoint associated with this approach to problems of mental retardation has been presented by Bijou (1966). Bijou points up rather sharply the disparity between the rapid advances in concepts and methods in various areas of psychology and the rather meager returns to date from psychological research on problems of mental retardation. Bijou feels that the main cause of the slow rate of progress is the traditional view of mental deficiency as a hypothetical condition of the organism inferred from behavior. He proposes that instead retarded development be viewed as principally a function of inadequate reinforcement and discriminative histories. Inadequate performance, whether in everyday life situations, intelligence tests, or school learning situations, should be taken simply as evidence that the appropriate behaviors have not been selectively strengthened by differential reinforcement in the past.

Analyzing the situations to which the retarded child is typically exposed, Bijou points out that a major cause of inadequate reinforcement histories may be the meager resources for selective reinforcement in institutions or foster homes. Further, the tendency to treat the retarded child as chronically ill, restricting his exploratory behaviors and not making reinforcements contingent upon differential behaviors on the child's part, might be expected to lead to a grossly restricted repertoire of discriminative behaviors.

The fact that the retarded child is often a behavior problem is likely to lead to considerable use of punishment on the part of those in charge of the child and it is well known that a reinforcement schedule depending largely upon punishment rather than reward is not conducive to building up repertoires of differential behaviors. Also, the effects of punishment are likely to generate side effects that would not be anticipated by the layman. For example punishing a child for socially undesirable verbal behavior without at the same time providing conditions suitable for differential reward of a more acceptable behavior may lead the child to avoid punishment by such artifices as speaking unintelligibly. Further, the situations in which punishment occurs and the people involved become conditioned stimuli for emotional responses and lead to avoidance, thus setting up conditions hardly conducive to constructive learning. Finally, unintentional reinforcement contingencies may lead to the specific shaping of socially undesirable behaviors. If, for example, the retardate has few behaviors available which obtain attention from adults, he may find by trial and error that this end is achieved only by socially undesirable acts and thus these would be unintentionally strengthened.

It must be recognized that the role of inadequate reinforcement histories in the causation of retarded intellectual functioning is documented by very little specific evidence. Indeed it is scarcely to be expected that Bijou's analysis can ever be fully verified at the human level by controlled experimentation. Though it is often easy to find evidence of extremely unfavorable conditions for learning in the background of the retarded child, it is never possible to determine whether the same inadequate learning condition in a child with different genetic background would have led to the same final intellectual performance. I suspect that the answer of the investigators in the operant-conditioning tradition would be that such comparisons are not essential. Proceeding on the premise that the analysis of retardation as summarized by Bijou is essentially correct, they would proceed to attempt to overcome specific varieties of retarded intellectual functioning by specific reinforcement procedures. Success in these efforts would certainly be taken as support for the theoretical analysis. However, the converse does not hold—failures to overcome these deficits in intellectual performance by training procedures would not justify the conclusion that the deficits are not attributable at least in part to unfavorable learning conditions in the past.

A considerable body of research on effects of early experience upon behavior in both animals and human beings provides some support for Bijou's analysis of the causes of mental retardation. A thorough review of this literature and its implications for problems of mental retardation has recently been presented by Haywood and Tapp (1966) and need not be duplicated here. It is clear that deficiencies in sensory input and in opportunities for varied learning experiences in early life may produce irreversible changes in adult performance in learning situations as well as emotional and motivational disturbances.

Further, studies providing comparisons of effects of different environments upon mental development of children leave little doubt that continued residence in custodial institutions where the children lack normal opportunities for learning experiences involving the continuing interaction with particular adults normally characteristic of a family often leads to a progressive decrease in measured intelligence and also to retardation in development of motor functioning and even physical growth. Some broad-scale studies (e.g., Gray & Klaus, 1964) have, by administering intensive programs of varied learning experience over periods of months to culturally deprived children, produced appreciable gains in measured intelligence, as compared to controls, and improved school learning performance. Thus, regardless of whether the analysis of retardation in terms of conditioning history can ever be supported in an extreme form, there is a tangible basis for anticipating major benefits from the kind of detailed analysis of deficient performance and individual programming of retraining procedures that are characteristic of the operant conditioning approach.

An excellent illustration of the operant conditioning approach in action in the context of mental retardation research has been presented by Sidman and Stoddard (1966). These investigators set themselves the task of determining whether severely retarded, institutionalized individuals could discriminate among simple geometric forms. Specifically, in the experiments to be discussed, to distinguish circles from ellipses. The type of individual with whom they proposed to establish communication and to make these determinations is typified by their first subject, a 40-year-old microcephalic idiot with an IQ below 20, with an active vocabulary limited to three words which he did not regularly use appropriately, and with the ability to respond only to a few simple commands such as "come here." Preliminary tests with methods of the sort ordinarily used to assess sensory capacities had suggested that this individual could not distinguish a circle from an ellipse, but the possibility remained that this failure was due to lack of a communication between the patient and the examiners.

Within the theoretical framework in which Sidman and Stoddard were working, it would be concluded that the individual could distinguish a circle from an ellipse if he could be trained to make different responses to stimuli of these two classes over a series of trials in which extraneous bases for differential response, such as positions of the stimuli, were eliminated. To arrive at this stage it would be necessary to shape the individual's behavior so that usable indicator responses would be emitted under the control of a reinforcement schedule, to arrange that the individual should be effectively stimulated by the two types of discriminative stimuli, and then to develop the desired contingency of responses on stimulus properties.

The first step in the training program for each subject was to find suitable reinforcing agents, that is, rewards. Even for the most severely retarded patients studied, this evidently was not difficult, with such measures as opportunities to play with young and enthusiastic assistants, the use of novel stimuli such as chimes, and the administration of novel food rewards such as small candies. Before behavior can be shaped in a particular direction, it must be emitted in considerable variety in the given situation. Thus the next step in training was always to permit the individual to habituate his original responses to the experimental situation, such as exploration, and then to strengthen approach behavior toward the situation by means of a variety of rewards. Then the learner was led by such devices as gestures, demonstration, and manual guidance to operate a simple manipulandum which operated a reward dispenser. The subject very quickly learned an easy discrimination. A panel of keys was presented to him, only one of which was lighted on any trial; operation of the lighted key led to reward and the dark keys to nonreward. This step could be accomplished almost without the occurrence of errors.

The remainder of the training program, by far the most difficult, was to find a sequence of steps whereby the dark keys could be changed gradually to lighted keys containing ellipses in such a way that the subject would continue responding only to the original lighted key which contained a circle. This differentiation procedure was accomplished by drawing very elongated ellipses on the dark keys and then very gradually over a series of trials increasing the illumination behind the dark keys so that the ellipses gradually became visible. The end result of this training regimen was a situation in which the subject responded to the lighted key with a circle, regardless of its position, even when surrounded with other lighted keys containing ellipses. Thus the investigators were able to conclude that the severely retarded patient's original apparent inability to distinguish circles from ellipses was due, not to a lack of sensory capacity, but to an inability to communicate appropriately with the examiners. More technically, in terms of operant conditioning theory, the patient originally differed from normal individuals in having no already established differential responses to the stimulus classes involved in the test; thus differential responses had to be established by suitable training before his discriminative capacity could be revealed.

Sidman and Stoddard present a series of extremely instructive accounts of their experiences in developing suitable training programs for a number of severely retarded individuals and also for very young normal children. Throughout, the emphasis is upon the prevention of errors, for one of their most ubiquitous observations is that the occurrence of errors tends to be highly disruptive, perhaps more so with retarded than with normal learners. Whenever in the sequence of steps during which an originally very easy discrimination is being transformed into the final more difficult one errors begin to occur, the result is nearly always a severe disruption of performance. Then the investigators must go back to a much earlier stage in the program and bring the learner up to the point of difficulty again with suitable revision, since continued training on the same stage, even with differential reward, generally does not get the learner past the point of difficulty.

Of particular theoretical interest is a major revision of a training program that these investigators were led to by an observation that might perhaps have been anticipated on the basis of discrimination theory (see Part A). A number of subjects had made rapid progress in the early stages of the program but broke down before reaching the final discrimination, and it occurred to the investigators that the concurrent changes in brightness of background of the incorrect keys and emergence of the ellipses on those keys might have been responsible. Over a series of trials, as the fading process proceeded, the ellipses became visible on the incorrect keys, but since the background brightness on those keys still differed considerably from that of the correct one, the subjects could continue to use the brightness cue even after the ellipse was visible. In terms of Zeaman and House's theory, the subjects were continuing to receive reinforcement for attending to an ultimately irrelevant dimension and thus were failing to learn to attend to the relevant dimension. When the brightnesses of correct and incorrect keys ultimately became equal, the subjects who had learned not to attend to the forms were, of course, unable to differentiate.

The remedy was a revised program in which the fading in of background brightness on the incorrect keys and the emergence of the ellipses were separated. Starting with a single correct key which was bright and contained a circle, positioned randomly among eight keys the remainder of which were all dark, the brightness of incorrect keys was gradually increased until it equaled that of the correct key, but with no forms present on the incorrect ones. Thus at the end of this stage the learner had to differentiate correct from incorrect keys by the presence of the circle on the former and thus was being rewarded for attending to the relevant stimulus dimension. After this stage had been successfully accomplished, faint ellipses began to appear on the incorrect keys and these were gradually filled in and darkened until the learner finally responded correctly to the lighted key containing the circle and avoided the lighted keys containing ellipses.

This study is especially significant in bringing out the importance of patient, detailed analysis of the steps required to proceed from the initial stage in which the learner's behavior is indiscriminate to a final one in which it is brought under the control of particular stimuli or stimulus dimensions, with errors at every stage in the training program being the principal basis for diagnosis, not of incapacities on the part of the learner, but of failures on the part of the training program. Whenever errors appear, it is assumed by these investigators that they represent a failure of the trainer to insure that the learner has been effectively rewarded for response to the appropriate stimuli or stimulus property at the preceding stage.

The long-term practical import of these demonstrations seems impressive indeed. If the severely retarded can be trained to discriminate circles from ellipses, there is little doubt but that they can be trained also to distinguish and respond appropriately to a variety of signs and symbols. Except for research purposes, it would not, of course, be feasible to tie up a large number of highly skilled individuals in the discrimination training of a single retardate. However, there are two important mitigating factors. Firstly, it is already apparent from the work on discrimination training that an effective training program, once developed, can be used for many different individuals with the amount of individual tailoring small relative to the effort involved in the original development of the program. Secondly, with the use of techniques that are rapidly being developed (see, e.g., Suppes *et al*., 1968) in computer-assisted education, it may soon be possible to implement rather large-scale training programs with relatively modest demands on skilled manpower.

XIII. *Contemporary Discrimination Learning Theories*

During the period when most of the learning theories discussed in preceding sections, and in particular those of Hull, Spence, and Skinner, were under development, discrimination learning was for the most part treated as subsidiary to the presumably more fundamental processes of conditioning and extinction. The prevailing assumption was that phenomena of discrimination could be interpreted simply by appropriate combination of such concepts as excitation and inhibition arising from studies of more elementary learning and conditioning.

More recently, however, the focus of research on both animal and human levels has shifted radically, discrimination becoming perhaps the central focus of both experimental and theoretical development. To be sure, current theories draw heavily upon concepts that have evolved from conditioning theory, but these concepts are augmented by ideas coming from other sources. As we have seen in the preliminary discussion of Part A of this volume, concepts having to do with attention and coding combine with principles of association and response selection in a multiple process conception of discrimination learning.

Of the specific formulations falling in this category, that of Zeaman and House (1963) has developed in closest interaction with research on mental development. Thus it will be convenient for our present purposes to begin with a rather detailed consideration of the model proposed by Zeaman and House and some associated lines of experimental investigation. In following sections I shall consider alternatives to the approach of Zeaman and House which are espoused by other investigators currently attempting to deal explicitly with problems of the development of learning abilities and also review some current research which seems especially relevant to these alternative conceptions. More extensive presentations of influential multiple process interpretations of discrimination learning are to be found in Lovejoy (1965, 1966), Neimark and Estes (1967), and Trabasso and Bower (1968).

1. background and experimental context of the attentional theory of Zeaman and House

Although it could presumably be generalized to other situations, the theory of discrimination learning proposed by Zeaman and House

(1963) has been developed in close connection with a lengthy series of researches in a particular type of discrimination experiment. Thus it will be convenient to present a résumé of their theory in terms of the associated experimental paradigm.

The apparatus has been adapted from the well-known WGTA used by Harlow (1959b) and his associates in their extensive studies of discrimination learning sets in monkeys. As adapted for use with children, the apparatus takes the form of a small desk in front of which the subject is seated. Immediately in front of the subject is a horizontal tray on which two stimulus objects are arranged in left-right positions, each being located above a well in which a bit of candy, which is the standard reward in these experiments, can be placed by the experimenter. The stimuli may be ordinary objects such as ink bottles or paper clips (used in so-called 'junk' discriminations) or, more often, plaques on which geometrical figures are drawn. Beyond the stimulus tray is a vertical panel containing a one-way mirror behind which the experimenter sits during the session. Lighting is arranged so that the experimenter can view the subject but the subject cannot see the experimenter. At the start of an experimental trial, the experimenter slides the stimulus tray back out of sight behind the screen, places the reward in one of the food wells and arranges the stimuli appropriately, then slides the tray out so that the subject can make a choice.

Suppose, for example, that the stimuli used in the experiment are a black circle and a black triangle and that the task is for the subject always to choose the triangle. The position of the triangle would be varied randomly between the left and right food wells so that, to master the task, the subject must respond on the basis of the stimulus property and not simply on the basis of position. Ordinarily a correction procedure is used, meaning that if the child lifts the correct object, he finds the reward under it at once, whereas if he lifts the incorrect object he is permitted then to lift the correct one and obtain the reward. To be credited with a correct response on any trial, the subject must select the correct object as his first choice.

A task involving simply discrimination between two objects is normally relatively easy even for severely retarded children and thus most of the studies conducted in this situation use problems involving at least two dimensions, one stimulus dimension being relevant and the other irrelevant to solution. Most commonly the two dimensions are form and color. A typical problem in this category might involve two forms, say triangle and circle, and two colors, say red and green. Four stimuli would be used, a red circle, a green circle, a red triangle, and a green triangle. On each trial two stimuli would be presented, each representing one value on each dimension. Suppose that in a particular case color were the relevant dimension with red correct and green incorrect. This would mean that over a series of trials the four stimuli would appear equally often, but on each trial one of the objects presented would be red and one green, and the reward would always be under the red stimulus regardless of whether it was a triangle or a circle.

Originally, Zeaman and House worked within the framework of the Hull-Spence type of association theory, which has been reviewed above. In terms of this kind of theory, the illustrative problem would involve four habits, each of which might acquire some habit strength over a series of training trials; the habits would be those relating the approach response to circular, triangular, red, and green stimuli. The response of approaching a red stimulus and receiving reward would occur on every trial, and consequently the strength of this habit would grow steadily along an exponential function to its maximum value over a series of training trials. The response of approaching a green stimulus would receive only delayed reward, and thus the maximum strength attainable by this habit would be lower than that of approach to red. The responses of approaching circular and triangular stimuli would sometimes receive immediate reward and sometimes delayed reward (equally often on the average) and consequently these would reach intermediate strengths. As the strengths of these habits diverged over training trials, the probability of choosing the correct stimulus would steadily increase to some limit which would depend upon the combination of motivation, reward, and stimulus differences. However, relatively early in their series of studies, Zeaman and House made some observations which suggested to them that this type of theory, a single-stage model as they termed it, was inadequate to represent discrimination learning.

In the classical literature on discrimination learning, the most common method of following the course of learning has been simply to plot the proportion of correct responses per trial or per block of trials for a group of subjects over the course of training. When Zeaman and House (1963) did this for a sizeable group of retarded children run on a simple discrimination in their apparatus, they obtained a learning curve of the usual form, starting near chance (50% in this case) and rising quite regularly along a negatively accelerated course to a final level of near 100% correct performance. When, however, they subdivided these subjects into more homogeneous subgroups on the basis of number of days required to meet the criterion of learning, they found that in general the subgroup learning curves did not have this form. Except for the most rapid learners, the subgroup typically remained virtually at chance for a considerable number of trial blocks and then rose relatively steeply to the final common level of nearly 100% correct responding. Thus it became clear that the group learning curve was not representative of the course of change in performance on the part of individual subjects.

As a consequence of this finding, Zeaman and House were led to reanalyze their data by means of "backward learning curves." This procedure, described by Hayes (1953) and by Suppes and Ginsberg (1963), takes as the reference point for each subject, not the initial trial of the experiment but rather the trial on which the subject meets the criterion of correct performance (20 correct responses in a daily block of 25 trials in the present instance). In preparing a backward curve, the protocols for individual subjects are brought together so that Trial 1 is the trial preceding the beginning of the criterion block for each subject, Trial 2 the trial before that, and so on; then the mean proportion of correct responses is computed for the trial preceding the criterion, the second trial preceding the criterion, and so on backward.

When this analysis was applied to the data for the group of retardates referred to above, the backward learning curves for the various subgroups were virtually superposed. All curves began with virtually horizontal segments over which proportions of correct responses were at chance, followed by a relatively sharp rise to 100% correct over about the last 40 trials preceding attainment of criterion. Thus the fast and slow learners were found to differ primarily in the number of trials before their learning functions began to depart from the chance level, and not in the slope of the final sharply rising portion of the curve representing the relatively abrupt transition from chance to criterion performance.

Zeaman and House concluded that a single-stage discrimination model of the type associated with the Hull-Spence theory did not fit the pattern of results revealed by these analyses. They felt that, if the mastery of a discrimination were simply a matter of the steady divergence of strengths of correct and incorrect habits as a function of training trials, there would be no basis for predicting the observed forms of the backward learning curves, with the long initial horizontal segments at chance followed by a steep rise to criterion, or the fact that subgroups of subjects with different learning rates differ primarily in the lengths of the initial plateaus.

The extent to which these analyses by Zeaman and House actually controvert a single-stage association theory of the discrimination process is difficult to say. There are several difficulties. Firstly, even if a single-stage theory necessarily implies that correct response probability should increase along a negatively accelerated growth curve as a function of trials, little is known concerning what precise form the backward curve should take if plotted for a group of subjects whose individual learning curves have differing slope parameters.

The only bit of evidence at all pertinent to this question that this reviewer has seen in the literature is an analysis presented by Trabasso (1963) in connection with a comparison of a single-stage, growth model with a two-stage, all-or-none model, the latter assuming that performance remains at chance until the final abrupt rise to criterion. In order to check out predictions from the growth model, Trabasso ran a group of computer-simulated subjects for such a model and plotted their mean backward learning curve together with the predicted backward curve for the all-or-none model (which is of course horizontal prior to criterion) and the curve for data obtained from human subjects. His observed data fit the all-or-none prediction closely, whereas the data for the simulated subjects learning according to a growth model yielded an average backward curve which rose in a negatively accelerated fashion over the trials preceding criterion. This fragment of evidence would tend to support Zeaman and House's reasoning. However, in fairness to adherents of the Hull–Spence system, it should be mentioned that Spence frequently emphasized that neither he nor Hull had worked out their theory in quantitative detail for the case of learning under a correction procedure which has been used in the series of studies by Zeaman and House.

In any event, the question whether the logic involved in Zeaman and House's original theoretical orientation was airtight is largely academic. The fact remains that their analyses of backward learning curves suggested that discrimination learning in their situation might represent a two-stage process, and they continued to conduct similar analyses with cumulative results which, over a period of time, began to provide rather impressive support for their position.

Backward learning functions on a color-form discrimination for groups of retarded children of differing MAs yielded an even sharper picture than the original analysis, with the different MA groups differing almost solely in the length of the initial plateau. In one analysis presented in their 1963 article, for example, groups of children with mean MAs of 2 years, 4 months, and 4 years, 6 months, were compared. The backward curve for the 4-6 group rose steeply from chance to near 100% performance over about 40 trials. The curve of the 2-4 group differed only in that it hovered around the chance level of 50% correct responding with no obvious trend for about 180 trials before beginning to rise, then like the 4-6 curve rose steeply to virtually 100% performance over about 40 trials.

Even more suggestive with regard to theory, perhaps, is the fact that similar patterns emerged when Zeaman and House compared performance for similar groups of children learning problems of differing difficulty. For example, in color-form discriminations, it is generally found, at least with retarded subjects, that color discriminations are considerably more difficult than form discriminations and discriminations based on two-dimensional objects. Backward curves for groups of children learning these various types of discriminations (Zeaman & House, 1963, p. 165), yielded a rather striking picture. The curves for the problems of varying difficulty differed greatly with respect to the length of the initial plateau over which performance remained at a chance level, but all were exceedingly similar in the steepness of the learning curve once it departed from chance and rose to criterion over an interval of 20-40 trials.

This finding was evidently one of the bits of evidence which suggested to Zeaman and House that the two stages in discrimination learning might represent an attentional process and a learning process, the initial horizontal segment of the backward learning curve representing a period before the subject had come to attend to the relevant stimulus dimensions for the given problem and the final steeply rising portion of the curve representing the process of associating the correct instrumental response with the relevant cues once these had been discovered.

2. a two-stage, attentional model for discrimination learning

The specific two-stage discrimination learning process to which Zeaman and House were led by their analyses was spelled out in the following assumptions (Zeaman and House, 1963).

1. On each trial of a discrimination experiment the stimulus display represents a variety of stimulus dimensions. The dimensions include

not only those specifically controlled by the experimenter, such as color and form in a color-form discrimination, but position and various other uncontrolled dimensions such as texture, distance, and the like, which are generally irrelevant to a discrimination problem.

2. The subject attends to only one dimension at a time.
3. The dimensional aspects of stimuli elicit attention; the result of the subject's attending is the perception of specific cues (for example, a circular form, the color red).
4. The specific cue which is perceived as a result of attending to a dimension elicits the instrumental response of approaching and lifting one of the stimulus objects.
5. Dimensions are broad classes of stimuli having common discriminative properties. Instances of dimensions which have been found to be important in discrimination learning by retarded children are form, color, size, position, and brightness.
6. If the subject's choice on a discrimination learning trial is rewarded, both the observing response and the instrumental response which he made on the trial increase in probability of occurrence. The term *observing response* refers to the behavior of attending to a stimulus dimension.
7. If the subject's choice on a trial is not rewarded, both his observing response and his instrumental response decrease in probability, and observing responses which do not occur on that trial receive complementary increases in probability. Similarly, when a response is rewarded, observing responses which did not occur on that trial decrease in probability.
8. In the two-alternative discriminative situation, the probability P of the overt response of choosing the correct stimulus object is given by the expression $P = ap + (1 - a)/2$, where a denotes the probability of observing the relevant dimension and p the probability of making the correct instrumental response given perception of the stimuli of that dimension. The probability of a correct instrumental response to an irrelevant cue is taken to be ½ under the usual experimental arrangements.

To elucidate these assumptions in terms of the standard discrimination-learning situation, let us take as an example the familiar color-form discrimination problem with color the relevant dimension, red being correct and green incorrect. It is assumed that at the beginning of the experiment a subject has some probability, generally low, of attending to the color dimension, some probability of attending to form, which is irrelevant in this problem, and some probability of attending to other uncontrolled and irrelevant dimensions. Since all irrelevant variables are carefully balanced over a group of subjects by the experimental design, we can assume further that at the beginning of the experiment, on the average, a subject's probability of choosing the correct object is ½, regardless of what dimension he attends to.

On early trials, the subject's choice of a stimulus object will be rewarded on about ½ of the trials and unrewarded on ½ in a random sequence, regardless of what dimensions he attends to. The probabilities of attending to the various dimensions will receive increments and decrements equally frequently on the average and thus will continue to hover around their original values. However, for each trial on which the subject happens to attend to the relevant dimension, color in this instance, his probability of making the correct choice response when attending to color will be increased (either because he makes the correct choice by chance and is rewarded or because he makes the incorrect choice and is not rewarded). Once the probability of making a correct choice when attending to the relevant dimension has risen above chance, so also will the probability of being rewarded when attending to the relevant dimension; the response of attending to color will begin to receive increments more often than decrements in probability and therefore will begin to rise above chance. Thus the process will accelerate like an autocatalytic chemical reaction, each increase in probability on the part of either the instrumental or the attentional response leading to an increase in the likelihood that both responses will gain in probability on subsequent trials.

According to this conception, the initial plateau of the backward learning curve, during which performance hovers around a chance level, represents a phase during which the subject is predominantly attending to irrelevant dimensions and thus is making no progress whatever toward mastery of the discrimination. When the subject finally happens to attend to the relevant dimension, the reward he receives will produce an increase in his probability of making the correct response when attending to that dimension, and the accelerative phase of his learning process will begin.

These assumptions concerning the subject's tendencies to attend to different dimensions and to make choices once he has perceived the cues of a given dimension are qualitatively similar in many respects to those of other two-stage theories of discrimination learning, notably those of Kendler and Kendler (1962) and of Sutherland (1959). Had Zeaman and House stopped at this point in their theory construction, their formulation would have suffered the same limitation as the others mentioned. The earlier theories have been influential in directing research in the hands of the originators, who have intuitive feel for the mode of operation of their only partially specified theories, but have been difficult for other investigators to understand fully or to test rigorously because of ambiguity concerning just what empirically testable propositions follow from the assumptions. However, Zeaman and House took the additional step of presenting their theory in a provisional quantitative form. At the cost of some simplification of their general notions and a temporary restriction to a specific type of experimental application, they made it possible for themselves and others to work out detailed implications of their assumptions.

The approach of Zeaman and House was to assume that each of the component learning processes included in their schema proceeds in accord with the same mathematical process assumed in a large class of currently influential models for simpler forms of conditioning and associative learning. The basic function prescribing the way in which the subject's probability of making a response changes upon occurrence of reward and nonreward was taken over from the probabilistic learning models of Bush and Mosteller (1955) and Estes (1950, 1959). In the interpretation of a discrimination learning situation involving a number of stimulus dimensions, the notation involved in the quantitative form of the theory becomes rather complex. Thus, as a basis for discussing the detailed properties of Zeaman and House's theory, it will be convenient to present the mathematical form of their theory for a particularly simplified case.

In the following résumé let us assume that we are dealing with a discrimination learning situation involving only two stimulus dimensions, one relevant and one irrelevant, so that the subject must attend to one or the other on every trial. Further, we shall assume that the subject's probability of making a correct choice when he attends to the irrelevant dimension is exactly ½ at the beginning of the experiment and remains at that value throughout; this would be only an approximation to the state of affairs generally assumed to obtain, but it gains a great deal in simplicity and will not materially affect any conclusions about the operation of the model. With these understandings, let us define the minimum necessary notation and recast Zeaman and House's assumptions about the discrimination learning process in quantitative form.

Definitions:

a_n	Probability of attending to the relevant stimulus dimension on Trial n
p_n	Probability of selecting the correct stimulus given that the subject attends to the relevant dimension on Trial n
P_n	Probability of a correct choice on Trial n

Assumptions:

1. Probability of a correct choice on any trial is related to the other basic quantities by the expression

$$P_n = a_n p_n + \tfrac{1}{2}(1 - a_n), \tag{1}$$

with the second term on the right expressing the assumption that the correct choice occurs only by chance when the subject does not attend to the relevant dimension.

2. If the subject attends to the relevant dimension on Trial n and his choice is rewarded, then

$$a_{n+1} = a_n + \phi(1 - a_n), \tag{2}$$

and

$$p_{n+1} = p_n + \theta(1 - p_n), \tag{3}$$

where ϕ and θ are constants with values between 0 and 1.

3. If the subject attends to the relevant dimension on Trial n and his choice is not rewarded, then

$$a_{n+1} = a_n - \phi a_n \tag{4}$$

and

$$p_{n+1} = p_n + \theta(1 - p_n). \tag{5}$$

4. If the subject attends to the irrelevant dimension on Trial n and his choice is rewarded, then

$$a_{n+1} = a_n - \phi a_n, \tag{6}$$

and

$$p_{n+1} = p_n. \tag{7}$$

5. If the subject attends to the irrelevant dimension on Trial n and his choice is not rewarded, then

$$a_{n+1} = a_n + \phi(1 - a_n). \tag{8}$$

Equations (2)-(4) make specific the notion that response tendencies are increased by reward and decreased by nonreward. In the former case, probability of the given response increases by a constant fraction (ϕ or θ) of the difference between its current value and unity. Equations (5), (6), and (8) simply express an assumption of symmetry together with the mathematical constraint that probabilities of mutually exclusive and exhaustive events must add to unity. To obtain Eq. (5), we note that the subject could be unrewarded when attending to the relevant dimension only if he made the incorrect choice response. In that event, the probability, $1 - p_n$, of the incorrect choice would be decreased by the factor $1 - \theta$,

$$1 - p_{n+1} = (1 - \theta)(1 - p_n),$$

which upon rearrangement is Eq. (5).

It is assumed that reward affects the probability of attending to an irrelevant dimension in the same way that it affects probability of attending to the relevant dimension. Thus when attending to the irrelevant dimension is rewarded, the probability, $1 - a_n$, of that response is increased by the fraction ϕ of the difference between its current value and unity, $1 - a_{n+1} = 1 - a_n + \phi[1 - (1 - a_n)]$; this last expression, rearranged, is Eq. (6).

For descriptive purposes, the constants θ and ϕ appearing in the basic learning functions may simply be regarded as rate parameters, the value of each obtaining for a given subject and problem being jointly determined by characteristics of the subject and characteristics of the task. From the standpoint of stimulus sampling theory (Estes, 1959; Neimark & Estes, 1967) these parameters would be interpreted as stimulus sampling probabilities. It would be assumed that there is available at the onset of each experimental trial a set of stimulus elements which are associated with the attentional responses. The parameter ϕ, then, represents the probability that any one of these elements is sampled (perceived) by the subject, and thus enters into the learning process, on any trial. Similarly, once the subject attends to the relevant dimension, a second set of ("discriminative") stimulus elements becomes available, and the parameter θ represents the probability that any one of these elements is sampled on a trial when the subject so attends.

The reader may note that there are two items missing from the list of basic learning functions, namely those representing changes in probability of the choice response when the subject attends to the irrelevant dimension. The reason for this omission is that in the simplified revision of the model presented here, it is assumed that probability of choice of either stimulus when the subject attends to the irrelevant dimension remains constant at ½ (whereas in the full model it would fluctuate randomly around an average value of ½).

For any experimental application, it is possible by means of Eq. (2)-(8), to derive expressions for the expected, or average, change in probability of the attentional response from any trial to the next. The procedure is to multiply the new value of this probability resulting from each possible outcome of the trial by the probability of that outcome and sum the quantities so obtained. Thus, with probability $a_n p_n$ the subject attends to the relevant dimension and chooses the correct stimulus object on Trial n, in which case the new probability of attending to the relevant dimension is, by Eq. (2), $a_n + \phi(1 - a_n)$. Proceeding similarly through all possibilities, one obtains

$$\begin{aligned} a_{n+1} &= a_n p_n[a_n + \phi(1 - a_n)] + a_n(1 - p_n)(1 - \phi)a_n \\ &\quad + (1 - a_n)\,\tfrac{1}{2}\,[a_n + \phi(1 - a_n)] + (1 - a_n)\,\tfrac{1}{2}\,(1 - \phi)a_n \\ &= a_n + \tfrac{1}{2}(1 - a_n) - \phi a_n(1 - p_n). \end{aligned} \tag{9}$$

To be precise, the term on the left-hand side of Eq. (9) should be altered to indicate that it represents an average value, but we shall take this to be understood and refrain from introducing any additional notation.

In the same way, we can take account of the changes consequent on all possible outcomes of Trial n and write an expression for the expected change in probability of choosing the correct stimulus given that the subject attends to the relevant dimension.

$$\begin{aligned} p_{n+1} &= a_n p_n[p_n + \theta(1-p_n)] \\ &\quad + a_n(1-p_n)\,[p_n + \theta(1-p_n)] + (1-a_n)p_n \\ &= p_n + \theta a_n(1-p_n). \end{aligned} \tag{10}$$

A comparable expression can be written for correct response probability on any trial as a function of the attention and choice probabilities on the preceding trial; but even in this simplified form of the model this recursion proves too cumbersome to be instructive. We can, however, gain some insight into the learning process represented in the model by examining the separate recursions for the attending and choice responses.

Considering first Eq. (9), we note that if $p_n = ½$, which would be the case on the average at the beginning of training on a new problem, the equation reduces to

$$\begin{aligned} a_{n+1} &= a_n + (\phi/2)\,(1-a_n) - (\phi/2)\,a_n \\ &= a_n + \phi(½ - a_n). \end{aligned}$$

If the average probability of attending to the relevant dimension were at chance, it would remain there; if not, it would change in that direction by a constant fraction of the difference between its current value and ½. In general, whether a_n increases or decreases on the average from its current value depends on the associated value of p_n, but regardless of the current value, the change will be positive if p_n is sufficiently large. In the limit, if $p_n = 1$, Eq. (9) reduces to

$$a_{n+1} = a_n + (\phi/2)\,(1-a_n)$$

and we see that a_n is driven toward unity.

Turning to Eq. (10), we note first of all that the average change in probability of a correct choice, given attention to the relevant dimension, is always positive, provided only that probability of attending to the relevant dimension is greater than zero. On every trial, the difference between the current mean value of p_n and unity is reduced by a fraction, the value of the fraction being directly related to a_n.

As long as a_n is initially greater than zero, the subject will be able ultimately to learn the discrimination. If initially a_n is small, then, although p_n will increase, on the average, from trial to trial from the very beginning of training, a_n will not necessarily increase on early trials; further, since the rate of growth of p_n will initially be commensurately small, a large number of trials may be required before the overall probability of correct responding [P_n in Eq. (1)] begins to increase. Once the curve for P_n begins to rise, however, the rate of climb will accelerate, since each further increment in either a_n or p_n increases the average rate of growth of the other.

Thus the model predicts the form of the backward learning curve observed by Zeaman and House, and, further, implies that the combination of a lengthy initial plateau with an ultimately steeply rising learning curve will result from a low initial value of a_n. It is for this reason that Zeaman and House concluded that differences in rate of discrimination learning between easy and hard problems for the same subjects, and between groups of subjects of differing mental age on the same problems, were, in each case, the result of differences in initial probabilities of attending to the relevant stimulus dimension.

Because the properties of this model are difficult to work out in detail analytically, Zeaman and House have made extensive use of computer simulation of their model in order to work out its implications for experimental phenomena. The general procedure is to write a computer program for a hypothetical individual who learns strictly in accord with the assumptions of the model and then to compute the performance of such hypothetical learners over a series of trials under a particular experimental arrangement. Since the model is basically probabilistic, these hypothetical subjects manifest individual differences in performance just like real subjects. Consequently it is necessary to "run" substantial samples of hypothetical subjects on the computer and to analyze the data in much the same way that one would analyze the data from a group of real children.

Examination of the performance of simulated subjects in the discrimination learning situation, reveals, first of all, that even when other parameters are held constant, performance is extremely sensitive to the initial probability of attending to the relevant stimulus dimension. For example, for groups of simulated subjects with the same learning rate parameter (θ value) and the same initial probability of the correct instrumental response, as the initial probability of attending to the relative dimension varies from .01 to .5, the learning function over 200 trials varies from an exceedingly flattened ogive which rises no more than 2 or 3 percentage points over the whole series to an exceedingly steep and sharply decelerated curve that passes 90% correct by about Trial 50. Similar variations in form and steepness of learning functions are observed when the number of irrelevant stimulus dimensions is varied while other parameters are held constant. Within any one of these families of curves, all of the functions rise more steeply as the θ values are increased, but rather large variations in θ are needed to produce modest effects.

An interesting, and so far as I know novel, implication of the model is the phenomenon which Zeaman and House term the "reversal midplateau." Suppose that in a form-color discrimination, the subjects during the original training series have learned always to choose the red rather than the green stimulus. If, now, the contingencies are changed so that the choice of green is rewarded, the learning function in terms of probability of choosing green is predicted to rise very rapidly from its initial value of near zero, reaching the chance level in a very few trials, then to level off for a few trials before again accelerating and proceeding toward a new asymptote of near 100%. The reason for the midplateau is that, from the outset of reversal training, the probability of the observing response for the formerly relevant dimension decreases on every trial, thus causing probability of a choice of that dimension to revert steadily toward chance.

If there are only two dimensions involved, one will observe no discontinuity, but if there are several irrelevant dimensions in the original problem, the observing response for the one which became relevant during reversal cannot begin to rise toward unity until the correct instrumental response associated with that dimension has received a number of reinforcements. Thus the model predicts that the midplateau should be accentuated as the number of dimensions involved in the problem increases and, other things equal, as the θ value for the instrumental response decreases.

The analysis of backward learning curves for simulated learners yields a pattern that we should expect from our examination of the simplified form of the model. Namely, variation in the initial probability of attending to the relevant dimension generates a family of curves which differ in the length of the initial chance-level plateau but which exhibit similar rates of rise to asymptote once they depart from the chance level.

One of the principal sources of evidence concerning Zeaman and House's discrimination theory has been experiments in which subjects learn an original discrimination and then are transferred to a new discrimination in which some of the stimulus-reward contingencies have been changed. A prediction of particular interest is that subjects should show positive transfer, that is facilitation of learning, on a relatively difficult discrimination if they have previously learned an easier discrimination involving the same stimulus dimensions. The reason is that the difficulty of the difficult discrimination must reside at least in part in a low initial probability of attending to the relevant stimulus dimension. If this attentional response could be learned during previous training on an easier discrimination, then the difficult discrimination would, in effect, be partially mastered before the subject had begun training on it and thus the performance curve for correct choices should be substantially accelerated.

House and Zeaman (1960b) report a pertinent experiment conducted with three groups of retarded children and a color-form discrimination. It had previously been found that for these children, as indeed for many other organisms, color-form discriminations are much more difficult when the stimulus display is a two-dimensional pattern than when it is a three-dimensional object. Thus they were able to use preliminary training on an object discrimination to produce relatively rapid learning of the response of attending to a relevant stimulus dimension and then to test transfer to a more difficult two-dimensional pattern discrimination. Three groups of children matched on MA were employed. The primary experimental group first learned a color-form object discrimination, for example, a three-dimensional red cross versus a three-dimensional blue triangle, and second, a two-dimensional pattern discrimination with the same stimulus dimensions involved, in the example a flat red cross versus a flat blue triangle. A control group was trained only on the second problem. A change-of-cues group also learned a three-dimensional problem followed by a two-dimensional problem, but in this case the cues representing each dimension were changed, for example, a green square versus a black circle in the first discrimination followed by a red cross versus a blue triangle in the second discrimination.

The control condition proved exceedingly difficult, only 4 of the 19 children being able to learn the discrimination within 500 trials, and the backward learning curve showed a very long plateau at the chance level of 50% correct responding before a final relatively steep rise toward asymptote. As predicted, the experimental group showed marked facilitation of learning on the difficult problem: 13 of the 19 children reached criterion on the first, three-dimensional discrimination, and all of these learned the subsequent two-dimensional discrimination within 20–30 trials. Of special interest is the change-of-cues group, which showed marked positive transfer, although not quite to the level of the primary experimental group. Two-phase theories which assume an initial phase of observing response learning, but with responses becoming associated with specific cues rather than with dimensions, would not, without special auxiliary assumptions, predict the substantial facilitation observed under the change-of-cues condition.

The theoretical picture is not entirely clear, however. Unfortunately, it must be noted that before full weight can be given to this favorable outcome, results will be needed for a third control group which is trained on a three-dimensional followed by a two-dimensional discrimination but with the relevant dimension, as distinguished from the positive cue, changed between the first problem and the second. In the absence of this control, the possibility must be considered that the positive transfer observed in the case of the change-of-cues group in the House and Zeaman study might have been based, not on learning of an attentional response, but on less specific factors. Examples of the latter are adjustment to the experimental situation and the previous experience of success with a problem generally similar to the final, difficult discrimination.

There seem to be no studies conducted especially to investigate the prediction of a reversal midplateau, but a few relevant data are reported by Zeaman and House (1963). Reversal-learning curves are shown for two groups of retardates who learned a position discrimination followed by a reversal; the two groups differed with respect to the amount of training in the first discrimination, and therefore presumably the probability of attending to the relevant dimension at the time when the reversal was initiated. In both cases the reversal curve shows some evidence of the midplateau, with the effect being considerably more accentuated for the group which had the lesser degree of training on the initial discrimination. On the basis of some simulation data, Zeaman and House suggest that not only the occurrence of the midplateau but the direction of the difference between the two groups was in accord with the predictions from their model.

The matter of the midplateau demands further investigation in a number of respects. First, it would be of much interest to know whether the same phenomenon can be predicted from other extant discrimination theories, as for example those of Atkinson (1958) and Lovejoy (1965). If not, then the phenomenon would be of special interest with regard to differentiating between the alternative theories. On the empirical side, the available fragments of evidence are scarcely a basis for any firm conclusions. More systematic studies are needed to determine the reliability of the effect and the conditions under which it occurs. With regard to the last point, it might be mentioned that similar plateaus have been observed under other conditions (Blehert, 1965). Since Blehert's experimental situation involved a choice among five response alternatives on each trial, rather than only two as in Zeaman and House's studies, it will be necessary to determine by suitable computer simulations whether Zeaman and House's model can predict plateaus in original discrimination learning under the conditions of Blehert's multiple-choice discrimination problem.

An aspect of discrimination learning which has recently received intensive study at a number of levels, retardates, normal children, monkeys, and rats, has to do with transfer of training as a function of relationships between the original discrimination problem and the transfer problem. Principal cases in addition to reversal learning are intradimensional and extradimensional shifts. The former denotes the case in which, following learning of a discrimination, a new problem is given involving the same relevant and irrelevant dimensions but with a new choice of cues on the relevant dimension. For example, if on an original color-form discrimination blue was the positive and yellow the negative cue, then on an intradimensional shift the subjects might be switched to a new problem with the same form cues irrelevant but with now green positive and red negative. The term "extradimensional shift" refers to a case in which the relevant cues in a new problem fall on a different dimension from that of the original problem. In the example, one would speak of an extradimensional shift if the subjects were shifted to a new problem with blue and yellow irrelevant, but with, say, circles positive and triangles negative.

Analyzed in terms of Zeaman and House's model, the task of the subject is quite different in different types of shifts. In the learning of an original discrimination, the subject must learn the observing response of attending to the relevant dimension and then must learn to associate the correct instrumental responses to the cues of the relevant dimension. Ordinarily the initial probability of the correct observing response is low and that of the correct instrumental response at chance (.5 in a two-choice discrimination problem), and both of these probabilities must be carried to unity before the problem is finally mastered. Similarly, following an extradimensional shift, the initial probability of attending to the now relevant dimension will be low and the probability of the correct instrumental response at chance, so, in general, it must be predicted that rate of learning an extradimensional shift should be about equal to that of learning the original discrimination. Following an intradimensional shift, however, the probability of the correct observing response will be high, since the subjects already have learned to attend to the relevant dimension which has not been changed, and learning of a new problem involves only carrying the probability of the correct instrumental response from .5 to unity.

Thus on an intradimensional shift learning should proceed more rapidly than in either of the other cases mentioned. Further, the initial plateau in the backward learning curve should be abbreviated and, if the final level of performance on the original discrimination is sufficiently high, the initial plateau of the intradimensional shift problem should be eliminated completely. On a reversal problem, again the initial probability of the relevant observing response will be high, but the probability of the correct instrumental response will be lower than chance. Thus reversal learning should proceed more slowly than an intradimensional shift, but with the initial plateau abbreviated as compared to original learning. Further, on the assumption that learning of an observing response usually proceeds much more slowly than learning of the correct instrumental response, reversal learning should generally be more rapid than original learning.

A more extensive discussion of investigations of the various types of transfer problems will be given below. At this point we might mention just one of the earlier relevant studies, an investigation by House and Zeaman (1962) which they analyzed in some detail in terms of the model. Three matched groups of retarded children were given original training on an object discrimination problem involving only one dimension, either color or form. Following original learning with either form or color relevant, subjects were transferred to a reversal, an intradimensional shift, or an extradimensional shift problem.

In good agreement with the qualitative analysis given above, learning of the second problem was very rapid for the intradimensional shift condition, only slightly less rapid for the reversal condition, and much slower, with a prolonged initial plateau near the chance level, for the extradimensional shift condition. Further, with plausible choices of parameter values, simulated data for stat children computed on the basis of the model yielded a set of backward learning curves very similar to those observed. It was noted also that learning of the extradimensional shift problems was somewhat more rapid than that of the original discriminations. House and Zeaman indicate that this finding might follow from their model given certain assumptions concerning relative strengths of competing irrelevant observing responses in the original and shift situations. However, before this aspect of the analysis can be properly evaluated, more adequate control data are needed to show whether some improvement might occur from one problem to another owing to nonspecific factors, as, for example, motivational level, understanding of instructions, or habituation to novel features of the experimental situation.

Although the successful application of the attentional model to these transfer problems has led to a confirmation of qualitative predictions which seems promising for the theory, considerable additional experimental analysis will be needed to determine the extent to which the theory can claim support. An important alternative interpretation of this last experiment which needs to be ruled out is that the positive transfer from original discrimination learning to the new problem following an intradimensional shift is a result, not of increased attention to the relevant dimension, but of adaptation to irrelevant stimuli which are present in both the original and transfer problems. As the experiment was done, the intradimensional shift and the reversal condition would evidently have an advantage over the extradimensional condition in this respect. The more recent literature on intra- and extradimensional transfer will be examined in a later section in order to determine how well confoundings of this sort, inevitably present in early experiments on a given topic, have been resolved.

Another application of the discrimination theory which Zeaman and House feel yields confirmatory evidence has to do with the introduction of a novel stimulus at a particular stage of discrimination learning. In one experiment, for example (House & Zeaman, 1958), the subject was presented with a single stimulus, which might be either rewarded or nonrewarded, on the first trial of an object discrimination series. Then, on the second and following trials, this stimulus was paired with another. The result was that regardless of the condition of reward or nonreward on the first trial, the subjects strongly preferred the new stimulus on Trial 2. This preference was interpreted as an instance of a familiar tendency on the part of both human and animal subjects to choose a novel over a familiar stimulus under a wide variety of conditions when other factors are approximately equal.

On the assumptions that novelty can be considered to be a discriminable property of stimuli, and that novelty enhances the probability of attention to the novel stimuli, Zeaman and House proposed that their model can handle this type of phenomenon. But, here we seem to run into a conceptual difficulty, for the theory of Zeaman and House differs from a number of other observing-response theories in assuming that the observing response relates to dimensions rather than to cues. One should observe effects of novelty only when new dimensions are introduced into a situation, not when new cues appear within dimensions already represented.

3. attention theory applied to mental retardation

A major implication drawn by Zeaman and House (1963) from application of their model to a variety of data from retardates and normal children is that the learning parameters (ϕ and θ) representing rates of learning of the attentional and instrumental responses are relatively independent of intelligence level, and, in particular, that even severely retarded children with IQs of the order of 25 are not distinguished by appreciably lower values than normals.

A strong interpretation of this conclusion would be that mental deficiency is something qualitatively quite different from the common-sense conception. Individuals of differing intelligence develop different habits of attention and stimulus selection. But they do not differ greatly with respect to learning ability in the sense of rate of formation of associations or retention of associations once formed. If correct, this view would seem to imply that such efforts as the search for chemical substances which may improve overall learning ability are doomed to failure.

Let us examine the factual side of the issue. In the context of the discrimination situation used by Zeaman and House, their conclusion is suggested quite strongly by analyses within the framework of their model. Their finding of markedly differing lengths of initial plateaus for groups of learners with different intelligence levels, together with similar final rates of approach to asymptote, is certainly the picture to be expected on the basis of their theory if initial tendencies to attend to relevant stimulus dimensions vary with intelligence, whereas θ values do not. The same is true of the analogous pattern found for discrimination problems differing in difficulty (Zeaman & House, 1963; Shepp & Zeaman, 1966). The extent to which these results should be taken as support for Zeaman and House's theory and of the implications following from it depends in part on the question of whether the same pattern of data could be predicted from other reasonable assumptions. This possibility requires careful analysis, which we defer to a later section.

If Zeaman and House's conclusion regarding the invariance of learning parameters is sound, one might expect that measures of intelligence would be relatively independent of rate of learning in situations which do not involve discriminations and thus where attention might be a less important factor. Whether there are any learning situations not involving discriminations is a moot point, but even if this state of affairs is an idealization, it should be approximated much more closely by simple conditioning than by discrimination experiments. The literature on measures of conditioning in relation to degree of mental retardation (see Denny, 1964) is fairly encouraging for Zeaman and House's position, for in general rates of conditioning have been found to vary little over a wide range of intelligence levels. Much the same appears to be true for simple motor-learning tasks, which again would be assumed to be less sensitive to attentional factors than discrimination or most types of verbal learning.

What is one to make of the overwhelming importance of attention in Zeaman and House's formulation? They suggest that individuals of differing intelligence levels vary primarily in their probabilities of attending to relevant stimulus dimensions at the start of any new problem. But it should be noted that in their model the probabilities of attending to various dimensions in any situation must sum to unity. Therefore one cannot, within their framework, meaningfully speak of a general trait of attentiveness comparable to such notions as degree of arousal or activation. From the first trial on any learning problem, the individual is always attending to some dimension or other. Therefore we must conclude that if the analysis is correct, the more intelligent individuals are those who are more likely to be attentive to the dimensions which are relevant in the kinds of problems used to test learning ability.

On the face of it, at least, this notion seems hard to accept. One logical implication is that even the most severely mentally retarded individual must be superior to normals on some types of problems.

Zeaman and House (1963, p. 217) cite a fragment of evidence implying that this might be so—a report by Weir and Stevenson (1959) that, among normal children, older children did more poorly than younger ones on a simple discrimination problem involving pictures of animals as stimuli. Instances of this sort are not very convincing, however, since it seems likely that the direction of the observed difference could readily have been reversed by a simple change in the instructions or in the way of explaining the task to the subjects.

A strong implication of Zeaman and House's interpretation would be that the remedy for mental retardation is generally to be found in modifying the environment rather than the individual. Any individual could be of normal or even superior intelligence if placed in an environment in which the stimulus dimensions to which he prefers to attend happen to be relevant for the solution of many types of problems. Even if the interpretation were not carried to quite such an extreme, the theory would nonetheless imply that training and education of retarded children should concentrate heavily upon ways of modifying attention to the relevant cues in various situations. One specific technique would be the extensive use of the easy-to-hard transfer paradigm, as for example by using three-dimensional stimuli in initial training of discriminations between letters, numbers, and the like. Another would be directing children's attention to relevant features of stimuli by instructions, by having them touch or trace outlines of forms, as the forms of letters during reading, and by arranging to associate rewards closely with relevant cues. Zeaman and House note that techniques of this sort have, in fact, been used by individuals (e.g., Fernald, 1943; Montessori, 1912; Strauss & Lehtinen, 1947) who report unusual success in remedial training of retarded or brain-injured children.

There can be no doubt that techniques that can attract the attention of the learner to relevant features of the task at hand are generally useful. Whether extensive reliance on such techniques could produce major changes in the ability of mentally retarded children to reach normal levels of performance in such areas as reading, calculating, and problem solving is quite another question. To develop a basis for even a tentative answer one needs firstly to evaluate the evidence for Zeaman and House's theory as it stands, and secondly to determine the extent to which it can usefully be extrapolated beyond the particular type of discrimination learning situation to which it has heretofore been applied in detail.

With regard to the first question, the data on easy-to-hard transfer fit well into Zeaman and House's analysis, but are certainly open to other interpretations. The similarity in the families of backward discrimination learning curves for discrimination problems differing in relevant dimensions (those for color being greatly retarded relative to form, etc.) and the families obtained for groups of children of different MAs on the same problems is suggestive. There do indeed seem to be grounds for hypothesizing that the rate at which the learner comes to attend to the relevant dimension of a problem is related in an important way to differences in difficulty among problems and differences in ability among learners.

However, the importance of the initial probabilities of attentional responses, as distinguished from differences in the parameters governing rate of learning of these responses, is so far as I know peculiar to Zeaman and House's theory and distinguishes their position from all other currently influential approaches to discrimination learning. Within their model differences in the rate of learning of relevant attentional responses do not produce the same kinds of differences in backward learning curves as differences in the initial probabilities of attentional responses. One wonders whether this feature of the theory is to be considered a minor idiosyncrasy, which could perhaps be eliminated by some revision in the assumptions that would not greatly change the implications of their model in other respects, or whether it is a necessary consequence of Zeaman and House's basic assumptions which could not be altered without major revision of the theory. Some theoretical analyses bearing on this point will be reported in a later section after we have completed a review of relevant empirical studies and have more systematically compared Zeaman and House's model with other discrimination theories.

4. discrimination shift studies

Because the consequences of various types of shifts between original discrimination and transfer problems are widely taken to be of a special diagnostic value with regard to types of discrimination theories, it seems useful to review this literature briefly. For a more extensive review with special attention to methodological problems the reader is referred to Shepp and Turrisi (1966).

Perhaps the most pervasive conclusion growing out of this literature is the consistent advantage in speed of discrimination learning on intradimensional as opposed to extradimensional shifts. This relationship seems well established for animals as well as for normal and retarded human subjects. Further, it seems quite well established that the advantage of the intradimensional shift increases with the degree of overtraining on the original problem. There is some indication that speed of learning an extradimensional shift, in contrast, decreases with degree of overtraining, but this evidence is not as consistent. Taken together, this set of findings is widely accepted as establishing the necessity for a two-stage discrimination theory involving either the learning of mediating responses or of selective attentional responses (see, e.g., Shepp & Turrisi, 1966, p. 117).

Let us first consider the status of evidence for the general conclusion. If learners, say, acquire a red-versus-green discrimination more rapidly following training on blue versus yellow than following training on circles versus triangles, what could be the basis for the facilitation? Experimental arrangements in this type of study assure that there could be no carry-over of specific stimulus-response associations from the first problem to the second in either case. It is conceivable that the tendency to ignore the irrelevant cues in the transfer problem is for some reason better established by previous training with the same relevant dimension, an assumption that would be required to interpret the studies in terms of Restle's adaptation-of-cues theory (1955). However, the advantage of intra- over extradimensional shift appears even when the irrelevant as well as the relevant cues in the new problem are new (Campione, Hyman, & Zeaman, 1965; Dickerson, 1967; Shepp & Eimas, 1964). There appears to remain no reasonable explanation other than that something which the subjects learn during the course of discrimination training, in addition to the correct stimulus-response relationships, carries over better to a new problem when the same relevant dimension is involved, and that this factor facilitates learning of the new discrimination. Three principal assumptions have appeared in the literature concerning the nature of this auxiliary learning process, and to these I would add a fourth.

In some respects the most parsimonious assumption is that the individual learns to observe, or attend to, the specific relevant cues in the course of learning any discrimination problem. This idea is usually credited to Spence (1936) and it was first given major emphasis in the discrimination theory of Wyckoff (1952) which, however, was never developed fully enough to yield specific interpretations of the various types of experiments we are considering. In order to account for the shift effects, the specific observing-response theories would have to be augmented by an assumption that transfer of these observing responses from one problem to another occurs more readily when the relevant cues fall on the same dimension than when they fall on different dimensions. In this line of thinking, the notion of a stimulus dimension might be taken simply as the experimenter's designation of a set of stimuli within which generalization among members is generally high as compared to generalization from these to stimuli outside the set.

The closely related, though logically distinct, approaches of Zeaman and House (1963), Lovejoy (1966), and, perhaps, Sutherland (1959) take the notion of a stimulus dimension as primary and assume that the first stage in mastery of a discrimination problem is to learn to attend to the relevant dimension. Then facilitation is naturally expected if the learners transfer to a problem involving the same relevant dimension, since they will be in a position to bypass the first stage of learning of the new problem. The notion of a stimulus dimension, or of an analyzer (Sutherland, 1959), has not been sufficiently fully explicated so that I, at least, can see what evidence might distinguish between two-stage discrimination theories of the specific observing-response type and those assuming attention to dimensions. Thus for present purposes there seems nothing to do but to group these together under one designation, say, two-stage attentional theories.

The most popular alternative to an attentional theory, among two-stage conceptions, has been the mediating-response hypothesis of Kendler and Kendler (1962). Again the task of critical analysis is handicapped by the fact that no one theory of this type has been spelled out in adequate detail. As I understand it, the Kendler's conception differs from that of Zeaman and House primarily in that, according to the mediational theory, not only does discrimination learning involve two stages, but also the final state of discrimination learning involves a chain of two or more responses. Rather than associating instrumental responses directly with the relevant cues in a situation, the learner is assumed first to associate some mediating response, presumably a verbal label in the case of human subjects and perhaps some more primitive counterpart in case of lower organisms, with the relevant cue and then to associate the correct instrumental response with the stimulus feedback from the mediating response. Thus, in a color–form discrimination with color relevant, the subject might learn to say to himself the word "red" whenever he sees a red stimulus, regardless of form, and then associate the correct instrumental response with the word "red."

Just as in the case of specific observing response theories, the mediating hypothesis has to be augmented by assumptions concerning generalization in order to handle the data of shift experiments. On the assumption that transfer of mediating responses is more likely to occur within than between stimulus dimensions, the hypothesis would account for the advantage of intradimensional over extradimensional shifts. Further there is an implication that phylogenetic and developmental stage of the organism would be of considerable importance since presumably adult human beings would be most likely to have suitable mediating responses available for many types of stimuli, at least to the extent that these involve language. The advantage of intradimensional and extradimensional shifts would be expected to increase as a function of age, paralleling the development of language skills, to be especially impaired in the mentally retarded, and to be less apparent in animals.

A conceptual difficulty with the mediational conception which has never been resolved for me by any published account has to do with the presumed function of the mediating responses. In the case of attentional theory, it is intuitively clear what selective attending would be expected to accomplish by way of facilitating discrimination. If the learner comes to ignore ambiguous cues in the situation, and to respond only to cues that are consistently related to reward and nonreward, he will necessarily perform more efficiently than if his behavior depends on all of the stimuli present regardless of their correlations with reward. But it is not obvious to me that developing a response chain consisting of a mediating response which leads in turn to an instrumental response would serve any especially useful function.

Consider, for example, a color-form discrimination in which red is positive, green negative, and circle and triangle vary randomly. An individual who learned in accord with the simplest form of one-stage association theory, would come over a series of trials to have the correct instrumental response associated with the color cues of a red triangle but correct and incorrect responses would be associated equally with the form cues. Consequently, if his response to the red triangle were a weighted average of these associative tendencies, he would not always respond correctly. An individual governed by a mediational model would presumably learn to make a mediating response, for example to say the word "red," whenever he saw a red triangle (or a red square); the correct instrumental response would then be associated with stimulus feedback from the word "red," and after the chain was established he would always respond correctly to a red triangle.

But, if the mediational assumption is simply embedded in an association theory, or added to the usual assumptions of Hull–Spence theory (as the Kendlers seem to have done), what is the basis for predicting that the individual will come to emit the mediating response "red" only in the presence of red stimuli? Why should this mediating response not be associated also with form cues just as would the instrumental response in a one-stage model? It appears that an attentional theory, or the equivalent, is required in order to account for the selective association of mediating responses with relevant cues, and thus that the mediational conception is not an alternative or a substitute for an attentional theory.

The other side of the coin is that the kinds of responses which one would expect to fill the role of mediators certainly do occur. Anyone who has watched a young child learning a color–form discrimination will have noted the frequency with which the child spontaneously squeals aloud "It's the red one," as he reaches for a red stimulus object. What is the function of these "labeling" responses?

A possibility, though most psychologists seem to discount it on intuitive grounds, is that they simply represent epiphenomena. A second possibility, and one for which a reasonable case can be made, is that the "labeling responses" do not mediate anything in the sense of serving as links of a chain leading from stimuli to instrumental responses but rather serve functions of facilitating memory storage and the direction of attention. If a learner, who on a given trial has perceived a relation between a certain stimulus property and the reward of an instrumental response based on it, has available a labeling response corresponding to the stimulus property, he will be in a position to rehearse this relationship following the trial and thus to facilitate recall of this relationship at the start of a later trial in which the same stimulus pattern occurs. Then, on the later trial, if the labeling response is recalled upon presentation of the stimulus, this response could serve to direct attention to the relevant cue, thus increasing the likelihood that the instrumental response would be based on this cue rather than an irrelevant one.

If this analysis is sound, it is fruitless to think of one-stage associational theories, attentional theories, and mediational theories as alternative models for discriminational learning and to attempt to discover from experimental results which is correct. Rather, we might better think of associative learning, selective attention, and mediating responses as different mechanisms, all of which can contribute to solution of discrimination problems and which may be present in various combinations in organisms at different developmental levels. Before much progress can be expected toward an overall theory, we need to determine empirically which mechanisms are operating in individuals at various developmental or ability levels in various types of learning situations.

It might be mentioned parenthetically that some types of discrimination situations, even for the higher organisms, seem to involve nothing more than associative mechanisms. I refer especially to discrimination in classical-conditioning experiments (see, e.g., the extensive review by Suboski, 1967). Even with college-student subjects, differential conditioning of the eyelid response with one stimulus reinforced and another stimulus not reinforced typically yields a picture of imperfect discrimination, with degree of discrimination as determined from the usual performance measures depending upon the degree of similarity between positive and negative stimuli. This fact could perhaps be accommodated within a theory of the type proposed by Zeaman and House, but apparently only at the cost of modifying some of the particular learning assumptions. Since within that framework, the development of selective attending depends upon reward contingencies, we could readily account for the absence of selective attention in classical conditioning experiments simply by assuming that the necessary reward contingencies are not available. However, if we interpret Zeaman and House's theory literally, conditioning would not occur at all in the absence of reward, so evidently one would have to assume that the conditioned response received some reward by virtue of its correlation with termination of the unconditioned stimulus on reinforced trials. If the attentional responses are not learned in the same situation, it must presumably be because they are not appropriately correlated with reward.

Such an interpretation would perhaps not be unreasonable. In the usual discrimination situation it is almost essential that the observing or attentional response continue during execution of the instrumental response (which usually consists in reaching for and lifting a stimulus object) and thus that it will be as closely correlated with reward as the instrumental response itself. In the classical conditioning situation the same would not be true, since occurrence of a response such as salivation or blinking of the eye does not depend upon, and in some cases may indeed be incompatible with, continued attention to some aspect of the conditioned stimulus.

In discrimination situations other than classical conditioning, the evidence seems quite uniformly to indicate that over a wide range of organisms including rats, monkeys, and both mentally retarded and normal human subjects, an attentional mechanism must be operative since intradimensional shifts are consistently easier than extradimensional shifts. It does not seem possible at present to be much more specific as to just what type of attentional mechanism is implicated, for the range of conditions that has been studied experimentally is severely limited. Most conspicuously, almost the entire relevant literature is limited to the visual discrimination problem, and almost exclusively to manipulations of a very few familiar dimensions such as form, color, brightness, and position. The one exception that has come to my attention is a study by Brown and Smith (1967), showing faster intradimensional than extradimensional shift for a tactual-kinetic discrimination.

The fact that the advantage of intradimensional over extradimensional shifts has been found quite consistently to increase with degree of overtraining on the original discrimination can readily be interpreted within Zeaman and House's theory, though it could scarcely be said to be predicted. If the response of attending to the relevant dimension is once overlearned to the point where it occurs with approximately probability 1 on every trial, further training would be wasted in the sense of contributing neither to performance on the given discrimination nor to positive transfer on new problems.[3]

It is quite possible, however, that if the criterion of mastery of an original discrimination is not too stringent it may be met at a point when the probability of the correct observing response is still somewhat short of unity. Zeaman and House, like Sutherland, assume that the rate of learning of the observing response is generally slower than that of the instrumental response, so one might expect that when a learner has just met a modest criterion of mastery of an initial discrimination, he would generally be in a state such that probability of the correct instrumental response to the positive cue would be near unity but the probability of the correct observing response would be somewhat less than unity. In that event, overtraining would produce an additional increase in probability of the correct observing response and thus also its probability of occurrence in a new problem involving the same relevant dimension.

---

[3] This property of the theory, however, is peculiar to the particular acquisition axioms that have been chosen. If, for example, the assumptions were modified slightly so as to incorporate the stimulus fluctuation process of stimulus sampling theory (see Estes, 1955), then continued training would increase the likelihood both that the observing response would be retained over a long period of time and that it would be transferred to new situations which share something in common with the old one.

A loose end with respect to interpretation of the intradimensional shift result is that several studies of discrimination learning in animals (e.g., Shepp & Eimas, 1964; Turrisi, Eimas, & Shepp, 1965) have indicated that the advantage of intradimensional over extradimensional shift may depend upon cues of the irrelevant dimension varying randomly with respect to their association with reward during original training. But in Zeaman and House's theory as presently constituted there is no intrinsic reason why this condition should be necessary. A possible basis for this finding, suggested by Zeaman and House (1963), is that when there is no variation of cues on the irrelevant dimension, those of the relevant dimension are distinguished not only by the property intended by the experimenter, for example, color, but also by novelty, since they are the only cues which change from trial to trial. If so, then an observing response which had been associated with the relevant dimension of one problem, might transfer to any cues in the new problem which were also distinguished by novelty.

In the case of intradimensional shift, if cues on the irrelevant dimension of the transfer problem vary from trial to trial these would tend to evoke the observing response and thus impair performance in comparison with the condition in which the correct observing response has been fully discriminated with respect to the distinguishing properties of the relevant dimension during original training.

It is probably not worthwhile attempting to push this explanation further or to evaluate it until further evidence is in hand indicating whether the variation of irrelevant cues during original learning is a similarly important condition in the case of human subjects. Also, it would seem reasonable to look for more direct experimental tests of the idea under conditions in which one might test for transfer where novelty is the only property common to an original and a transfer problem which might serve as the basis for positive transfer.

A large portion of the literature on transfer of discriminations has been concerned with comparisons between reversals and other types of shifts. This emphasis seems to be based on other than logical considerations, for the reversal situation is particularly difficult from a theoretical viewpoint and not amenable to a satisfactory analysis in terms of any extant theory. In the case of intradimensional shift, presumably the only possibility of transfer is a carry-over of the observing or attentional response to the new situation. There can be no positive or negative transfer of the instrumental response, since the cues to which it has become associated in the first problem are not present in the second. On an extradimensional shift, possibilities are either negative transfer or no transfer depending upon particular conditions. In the common case when the relevant and irrelevant dimensions are simply interchanged, negative transfer would be expected, since the observing response acquired during original learning would be evoked by irrelevant stimuli on the transfer task. The cues to which the instrumental response had been learned would also be present in the transfer problem, but now uncorrelated with reward so they would initially yield correct responding only at a chance level.

If all of the cues of the original discrimination are replaced with new ones in an extradimensional shift, then there is no basis in the attentional theory to predict any difference in speed of discrimination learning as compared to learning of the original problem.

In the case of reversal, both positive and negative transfer should be expected and thus the net outcome must depend on the balance of the two. Since the assignments of reward and nonreward to the formerly positive and negative cues of the relevant dimension are simply interchanged on reversal, the correct instrumental response will normally have zero or near zero probability at the beginning of reversal learning, and this factor by itself would lead to an increase in error rate as compared to early trials on the original problem. Since this was the only form of transfer envisaged in classical association interference theory, the earliest studies showing faster learning under reversal than under original learning conditions were quite surprising.

In terms of the attentional theory, we should of course expect the correct observing response to be carried over from the original to the reversal problem since the relevant dimension is the same, and this factor by itself would produce positive transfer. In general there is no basis for predicting which of the two opposed sources of transfer should dominate in any situation, and by suitable choices of parameter values in the attentional theory either positive, negative, or no transfer can be predicted.

In the study of House and Zeaman (1962) discussed above, speeds of learning under reversal and intradimensional shift proved approximately equal, and both were superior to that under extradimensional shift. A similar result was obtained by Eimas (1967) with normal children, the same relationships holding for both overtrained and nonovertrained groups. More often, however, overtraining is found to facilitate reversal more than other types of shift. For example, Furth and Youniss (1964) found for normal children that, in a nonovertrained group, intradimensional shift was superior to either reversal or extradimensional shift, the latter two being approximately equal, whereas in an overtrained group intradimensional shift and reversal were both superior to extradimensional shift. A study by Campione *et al.* (1965) fits the same general picture.

These findings can be accommodated within the attentional theory on the assumption that the learning-rate parameter for the instrumental response increases in value as a function of overtraining. There appears to be no explanation at hand as to why this relation should obtain. The overtraining results recall the "overlearning reversal effect" that has been the subject of much analysis in the recent literature on animal discrimination learning (for reviews see Sperling, 1965a, b). Increases in speed of learning with successive reversals are in some instances quite striking, even the rat being able to reach the stage at which it is able to reverse a discrimination in a single trial. However, the overlearning reversal effect has been found consistently only for positional discriminations, and there the presence or absence of the effect has been found to depend upon reward magnitude (Theios and Blosser, 1965). In discriminations based on visual cues, the phenomenon comes and goes in a perplexing fashion, usually appearing in, for example, studies conducted by Mackintosh (1965) with rats, but uniformly failing to appear in similar experiments conducted by Warren and his associates (Beck, Warren, & Sterner, 1966; Warren & Warren, 1967) with cats, and so on.

A major possibility to be considered is that special susceptibility of reversal learning to the effects of overtraining has to do with factors that are in a sense extraneous, that is, that one should not expect to be incorporated into a discrimination theory. One possible factor, suggested by D'Amato and Jagoda (1961) is that during prolonged overtraining the subject has no experience with the negative stimulus, especially in the positional discrimination, so that some forgetting of the consequences of the incorrect instrumental response may occur. The result would be that after much overtraining the individual's tendency to avoid the negative cue would be reduced and thus there would be less negative transfer to the acquisition of the instrumental response of approaching the formerly negative cue in the reversal problem.

Similarly, the possibility comes to mind that during the long series of correct responses during overtraining there may be some habituatory decrement which would tend to increase the tendency to spontaneous reversal. This notion would fit in particularly well with the pattern of results in the rat, since it is well known that that animal tends spontaneously to reverse a positional response following a consistent sequence of rewards whereas the same "alternation tendency" is less prominent in the case of response to cues which vary in position.

The finding that improvement in reversal learning with overtraining occurs more regularly with human subjects than with animals, and more with older than with younger human subjects, has been taken by the Kendler's (1962) as support for a mediational theory, since the most clearly documented instances of mediating responses are verbal, and verbalization in children increases on the average as a function of age. The possibility needs to be ruled out empirically, however, that older children are also more susceptible to habituation (or, to put it another way, boredom) than younger children and for that reason are more ready to reverse instrumental responses.

A number of other comparisons of reversal with other types of shifts in different types of learners can be taken to support the Kendler's interpretation in terms of developmental stages of mediating responses, but at the same time do not exclude the other type of interpretation in terms of a habituatory process. For example, Kendler and Kendler (1959) found that within the same group of children faster learners showed an advantage for reversal over nonreversal shift, whereas the converse was true for the slower learners. (Here nonreversal shift involves making the previously relevant dimension constant in the new problem and the previously varying, irrelevant dimension, relevant). Several other studies have shown the advantage of reversal over nonreversal shift to increase with overtraining within groups of children of the same age. Sanders, Ross, and Heal (1965) found no differences between the reversal and extradimensional shift for retarded children but a difference in favor of reversals for normal children, and other studies have shown that for normal children the advantage of reversal over extradimensional shift increases with age.

The difficulty in evaluating alternative interpretations of the effects of overtraining on reversal is due in part to the very limited class of experimental paradigms that have been used, in particular to the strict limitation to two-choice problems. By using problems with more than two stimulus alternatives it would be possible to study separately the effects of overtraining, unlearning, or extinction of the previously correct instrumental response as distinguished from the acquisition of the one which is correct on the "reversal" problem.

5. general comments on discrimination theories and related experimental analyses

There would seem to be no doubt but that the formulation of discrimination theories involving attentional and mediational concepts and the related lines of research on discrimination learning and transfer in normal and retarded children are developments of major importance from both practical and theoretical standpoints. The most hopeful auger for substantial advances in the training of the mentally retarded is the fact that for the first time we are beginning to see instances in which theoretical analyses lead to changes in training conditions or procedures, often apparently quite modest in nature, which have very large effects on the course of learning. One striking case in point is the effect of object-to-pattern transfer (House and Zeaman, 1960b) which we have already discussed; another is the correlation of qualitatively different rewards with different cues of a discrimination (Shepp, 1963).

Conversely, procedures which from the viewpoint of an attentional theory would be expected to be deleterious have been found in some cases to exert disastrous effects on the ability of retarded children to learn discrimination problems. One instance reported by Zeaman and House (1962) is that of subjects who had a prolonged experience of failure on an overly difficult discrimination problem (which from the viewpoint of the theory might be expected to lead to the extinction of attentional responses) and then were found unable to learn simple problems which had previously been well within their capacities. In a recent study by Touchette (1968), which was evidently not conducted explicitly within the framework of an attentional theory though it might well have been, a group of retarded children who were led by a carefully programmed sequence of gradations in the stimulus displays to attend to the relevant cue of a relatively difficult visual discrimination all mastered the discrimination, whereas nearly all of the group given conventional trial-and-error training on the same problem made no measurable progress whatever.

At the theoretical level, it seems to me quite beside the point to devote much energy at the present stage to such matters as trying to decide between different extant theories such as the attentional theory of Zeaman and House and the mediational formulation of the Kendlers. Both of these formulations are proving extremely useful aids to research, but as I have indicated above, it is probably best at this point to think not of alternative theories, but rather of a number of different mechanisms, associative, attentional, mediational, all of which may have places in a viable theory of discrimination learning.

The attentional formulation of Zeaman and House is a particularly apt starting point for further theoretical efforts in this area at the moment because they have gone further than others in spelling out a specific set of assumptions which can be brought to quantitative test. In the immediate future, considerable purely theoretical work needs to be done by way of changing particular assumptions or subsets of assumptions in their theory and working out the consequences of these by computer-simulation methods much as Zeaman and House have done for their own initial formulation. By this means it may be possible to gain insight as to which features of their theory are relatively robust and which features can readily be altered in order to circumvent some of the difficulties and apparent paradoxes that have arisen in interactions with research to date.

Two basic issues with regard to Zeaman and House's formulation are in particularly urgent need of clarification. The first is the question whether the first stage of the two-stage process is correctly identified as one involving observing responses, or attention as the concept is ordinarily conceived, as distinguished from a process of coding or selective memory storage. The second issue, presuming that further analysis continues to implicate an attentional mechanism of the general type conceived by Zeaman and House, is whether the present assumption of a closed set of attentional responses is correct.

Some of the difficulties and apparent paradoxes that have risen in applications of Zeaman and House's model, and which have been met with various *ad hoc* suggestions, are intimately tied up with the assumption that the individual is always attending to some dimension in any situation, the consequences of learning operations being simply to shift attention from one dimension to another. An alternative is to assume that at any time an individual may be attending to some or to none of the stimulus dimensions represented in a situation. On the latter type of assumption, learners of higher developmental levels could be assumed, as seems intuitively natural, to have higher probabilities of attending to all of the dimensions of an experimental situation than those of lower level, and this difference could facilitate rather than hinder discrimination learning on the part of the "brighter" individuals.

XIV. *General Comments on Methodology*

The general impression created by this review of empirical and theoretical literature is that a discipline having to do with the development of learning ability has only begun to take form. Some of the essential ingredients are beginning to appear, but major efforts up to the present have been concentrated upon a few facets of a problem which needs broad-scale and sustained attack.

It is tempting to seek grounds for optimism in the impressive quantitative increase in research on learning abilities and disabilities during the past few years. Faith in the idea that the urgent practical problems of this area will be solved by sheer industry must be tempered, however, by observation of the rapidity with which diminishing returns set in as the volume of work using the same methods and concepts is expanded. The analyses set forth in this volume almost without exception point to the conclusion that major advances at both the conceptual and practical levels come from close interactions of theory and experiment.

Perhaps one of the most pervasive conclusions arising from an examination of research within the framework of contemporary learning theory is the impossibility of sharply dissociating learning abilities from the results of learning. In every type of learning situation which has been subjected to considerable experimental study one finds that, when an individual of a given age is exposed to a learning task, both what he learns and how he learns it depend in major ways upon the results of past learning in other situations. Though relevant empirical evidence is meager, there are some theoretical reasons for assuming that individual differences in measured rates of learning must depend in part upon individual differences in elementary capacities which are to a major extent genetically determined. There are equally cogent theoretical grounds, with much more substantial empirical support, for believing that at least in the case of human learners, individual differences in ability at any age also depend strongly upon cumulative effects of past learning as embodied in learning sets and strategies.

This conclusion leads one to envisage two main tasks for those currently interested in advancing a science of learning ability.

1. The first of these, logically at least, involves the identification of elementary capacities and the tracing of the course of their maturation in the developing organism. Quite possibly, major theoretical progress on this line will have to wait upon the advance of behavioral and neurophysiological analyses of simple forms of learning and related behavior-genetic analysis. At a more molar level, the comparative approach to the study of mental retardation espoused especially by Ellis (1969) is presumably directed at least in part toward locating basic capacities by comparing the learning of normal and retarded individuals in standard laboratory situations.

Although there seems to be no satisfactory general rationale for the comparative approach, it is possible to discern some glimmers of success for a few particular lines of research pursued within this framework which are influenced also by developments in learning theory. Perhaps the most important case in point is the short-term memory process revealed by converging lines of investigation, some arising within mathematical learning theory (e.g., Atkinson & Shiffrin, 1968) and some within the comparative approach to mental retardation (Ellis, 1969). It must be emphasized that in both cases this process has been isolated only as a result of intensive experimental and quantitative analysis. For those with major interest in understanding individual differences, it may be a bit of a disappointment to find that short-term memory capacity appears to be surprisingly invariant over a wide range of ages and IQ levels (Calfee, 1969; Ellis, 1970). However, the isolation of this relatively constant factor has been a useful step toward uncovering the operation of a less homogeneous set of dispositions having to do with rehearsal and recoding strategies which do differ widely among individuals of differing overall intellectual levels.

2. The second main task is identifying learning processes whose products are important determiners of subsequent learning, investigating the conditions of initiation and maintenance of these processes, and tracing out the mode of their organization at various stages of development. Of the most widely employed conventional approaches to the study of learning ability, neither the comparative nor the longitudinal is likely to prove adequate taken by itself. Just as the comparative approach to the identification of capacities appears of limited potentiality unless pursued in conjunction with neurophysiological and genetic methods, so also the conventional longitudinal approach appears to offer little promise of substantial progress toward unraveling the complex network of dependencies of present upon past learning. In

this instance the necessary auxiliary tools will presumably have to be drawn from concurrently developing theories of learning and cognitive processes.

Some of the considerations involved are well illustrated in a recent work of Kendler, Kendler, and Marken (1969) on developmental trends in the use of representational responses by children in dealing with simple discrimination and concept learning problems. The procedure which appeared to be the principal reliance of many earlier investigators in this area, namely the amassing of data for performance on the same standardized task by children at different age levels, has proven all but barren for theoretical purposes. One can find much more promise for the future in the studies, so far few in number, which compare, at different age levels, either relationships between performances on different tasks or processes inferred from performance within a theoretical framework. As illustrated in the Kendlers' work, this latter type of investigation will often require the devising of novel experimental procedures dictated by theoretical considerations (e.g., the half-reversal shift in the 1969 study). Frequently there will also be required some modification of the basic learning task as it is applied at different age levels in order to isolate changes in performance attributable to the process under investigation from those attributable to other processes which are correlated in the course of normal development.

To what extent can one assume that broadly conceived programs of theoretical and experimental attack upon the general problems of learning ability and mental development will solve the practical problems of dealing with the mentally retarded and other individuals with more special learning disabilities? In attempting to come to grips with this enormously complex question, I have found it useful to think in terms of the schema proposed by Zigler (1967, 1969) who distinguishes between "pathological" and "normal" cases of low intelligence or learning disability. If it be granted that performance on any learning task is the result of some complex assemblage of genetic and experiential determiners, then one must expect to find in any population of individuals a more-or-less normal distribution of scores no matter how learning ability is measured. The individuals who fall toward the bottom of the distribution may pose special problems for society, but from a theoretical standpoint they differ only quantitatively from the individuals we term "normal"or "superior."

The understanding of the individual differences in manifest learning ability that arise throughout the range of this normal distribution must be part and parcel of the general problem of understanding the organization of learning processes in the individual and the way in which this organization develops as a function of specific kinds of experience. Further, I think that any substantial advances in improving the learning performance of individuals who fall at the lower end of this distribution must be expected to come from general theoretical advances, rather than from comparisons of individuals falling at the middle and lower portions of the distribution on various particular tasks. Such comparisons have their uses, as we have observed, but I have been able to find no rational basis for the heavy emphasis on comparisons between "normals" and "retarded" rather than between "normals" and "accelerated" individuals. In fact there are reasons to think that the latter type of comparison might often be the more fruitful, since it is generally more advantageous to study a trait or process by finding cases in which it is exemplified to an unusual degree than by concentrating on cases in which it is poorly represented or absent.

The residual class of learning disabilities in Zigler's schema is that including individuals who exhibit identifiable organic defects or who have suffered from specific abnormalities of experience, for example extreme social deprivation in early childhood. These pathological cases must necessarily be so heterogeneous in character that one can scarcely expect to accomplish much in attempting to deal with them on general theoretical principles. At a research level, comparative studies must be important, but it would seem that the comparisons should be concentrated, not on differences between these individuals and normals, but rather on differences between the performance of these individuals and that required for social adjustment. The focus of attention needs to be, not on the locus of the disability of a particular individual, but rather upon his remaining resources which might be capitalized upon to compensate for the particular pathology.

Of the various schools of investigators engaged in research upon learning in the mentally retarded, those associated with the operant conditioning tradition seem to take most seriously the strategy of analyzing the performance of the retarded individual, comparing it in detail with that desired, and generating specific research programs calculated to yield information as to how to attain the desired resultant. At the present time, the specific methods of operant conditioning are applicable largely to problems that arise with the severely retarded, in whom it is a major accomplishment to teach the appropriate response to a simple command or to develop the rudimentary discriminative behavior needed to permit adequate assessment of sensory capacities.

Though some of the accomplishments obtained by application of operant conditioning methods to the severely retarded seem quite useful, indeed impressive, perhaps even greater fruits could be hoped for by application of the same general strategy at the level of the educable retarded. The vast preponderance of learning research at this level seems to have been concerned primarily with comparisons between retarded individuals and normals in various simplified laboratory learning tasks. It is very difficult to find cases in which investigators have analyzed the deficient performance of the higher-grade defective in school learning or other situations involving cognitive or verbal functioning, proposed a specific hypothesis as to the deficiencies in the previous learning history, and developed individual training programs which, on the basis of the best available learning theory, might be calculated to lead to the desired performance.

With even the meager beginnings of a theory of cognitive development now available, it seems clear that one could prescribe an environment and a sequence of learning experiences for a normal child which would lead him in adulthood to exhibit discriminative and problem-solving behaviors of the kind which characterize the mentally retarded. Less is known as to the type of environment which might be expected to lead any except perhaps the severely brain-damaged child to develop the learning strategies which now characterize normal adults, and still less as to the conditions, if any, which might produce this result in the case of the mentally retarded adult. However, so much is at stake from both theoretical and practical viewpoints, that major research efforts directed toward these possibilities appear to be called for.

It must be emphasized in this connection that the amounts of experience involved in the development of learning strategies in the normal child must be very large compared to those provided in ordinary laboratory experiments. Despite the formidable practical difficulties involved, high priority must attach to the need for conducting sufficiently long-term experimental studies to provide substantial evidence concerning the reversibility of the deficits in discriminative learning and cognitive abilities which characterize the mentally retarded.

With respect to the principal motif of the present review, one must conclude that, on the whole, contacts between learning theory and the empirical study of mental development have been sparse and unsystematic. Efforts toward more effective interplay have so far appeared almost exclusively in connection with mental deficiency and learning disability. Here the results have been quite encouraging, with particularly fruitful indications of progress in both theory and application arising from sustained and intensive theoretical-experimental analyses of short-term memory, discrimination learning, and organization and retrieval strategies in verbal recall.

Reasons for the relative isolation of the two disciplines in the past are evidently two sided. The study of mental development has only recently begun to shift from primarily observational to experimental methods. On the other hand, learning theory, in this country at least, has been closely associated from the outset with laboratory studies of learning. Further, for reasons which surely have nothing to do with any broad strategic considerations, the subjects of laboratory analyses of learning were for a long period confined largely to human adults and a few species of lower animals. Naturally enough, the concepts and principles of contemporary learning theories reflect these limitations and provide scant resources for expressing ideas regarding developmental processes.

That this theoretical one-sidedness has not been felt keenly by most investigators of learning perhaps owes largely to the heavy concentration of research on relatively simple forms of conditioning and learning, where similarities over age levels are more conspicuous than differences. It is well known that diagnosing a weakness in scientific strategies rarely leads in any direct way to reform, but even so there are some grounds for guarded optimism in the present instance. At least a partial remedy may already be taking form in the increasing emphasis on more complex phenomena of learning and memory, with the attendant increasing need to deal with the organization of learning and cognitive processes.

References

Achenbach, T., & Zigler, E. Cue-learning and problem-learning strategies in normal and retarded children. *Child Development*, 1968, **39**, 827-848.

Allen, G. A., Mahler, W. A., & Estes, W. K. Effects of recall tests on long term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 435-447.

Asratyan, E. The functional architecture of instrumental conditioned reflexes. *Conditional Reflex*, 1967, **2**, 258-272.

Astrup, C., Sersen, E. A., & Wortis, J. Conditional reflex studies in mental retardation: a review. *American Journal of Mental Deficiency*, 1967, **71**, 513-530.

Atkinson, R. C. A Markov model for discrimination learning. *Psychometrika*, 1958, **23**, 309-322.

Atkinson, R. C., & Shiffrin, R. M. Human memory: a proposed system and its control processes. In J. T. Spence & G. Bower (Eds.), *The Psychology of Learning and Motivation*, 1968, **2**, 89-195.

Bandura, A. Vicarious processes in a case of no-trial learning. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 1965, **2**, 3-55.

Bandura, A., Grusec, J. E., & Menlove, F. L. Some social determinants of self-modeling reinforcement systems. *Journal of Personality and Social Psychology*, 1967, **5**, 449-455.

Barnett, C. D. Stimulus generalizations in normals and retardates on a visual-spatial task requiring a voluntary response. Unpublished doctoral dissertation: George Peabody College, 1958.

Barnett, C. D., & Cantor, G. N. Discrimination set in mental defectives. *American Journal of Mental Deficiency*, 1957, **62**, 334-337.

Barnett, C. D., Ellis, N. R., & Pryer, M. W. Stimulus pretraining and the delayed reaction in defectives. *American Journal of Mental Deficiency*, 1959, **64**, 104-111.

Barnett, C. D., Ellis, N. R., & Pryer, M. W. Serial position effects in superior and retarded subjects. *Psychological Reports*, 1960, **7**, 111-113.

Barrett, B. H., & Lindsley, O. R. Deficits in acquisition of operant discrimination and differentiation shown by institutionalized retarded children. *American Journal of Mental Deficiency*, 1962, **67**, 424-436.

Bauer, F. J., & Lawrence, D. H. Influence of similarity of choice-point and goal cues on discrimination learning. *Journal of Comparative and Physiological Psychology*, 1953, **46**, 241-248.

Beck, C. H., Warren, J. M., & Sterner, R. Overtraining and reversal learning by cats and rhesus monkeys. *Journal of Comparative and Physiological Psychology*, 1966, **62**, 332-335.

Belmont, J. M. Perceptual short-term memory in children, retardates, and adults. *Journal of Experimental Child Psychology*, 1967, **5**, 114-122.

Bensberg, G. J. *Teaching the mentally retarded*. Atlanta, Georgia: Southern Regional Education Board, 1965.

Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology*. 1965, **2**.

Berkson, G., Hermelin, B., & O'Connor, N. Physiological responses of normals and institutionalized mental defectives to repeated stimuli. *Journal of Mental Deficiency Research*, 1961, **5**, 30-39.

Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. *Journal of Personality*, 1961, **29**, 303-320.

Bialer, I., & Cromwell, R. L. Task repetitions in mental defectives as a function of chronological and mental age. *American Journal of Mental Deficiency*, 1960, **65**, 265-268.

Bijou, S. W. Patterns of reinforcement and resistance to extinction in young children. *Child Development*, 1957, **28**, 47-54.

Bijou, S. W. Operant extinction after fixed-interval schedules with young children. *Journal of the Experimental Analysis of Behavior*, 1958, **1**, 25-29.

Bijou, S. W. Discrimination performance as a baseline for individual analysis of young children. *Child Development*, 1961, **32**, 163-170.

Bijou, S. W. A functional analysis of retarded development. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*. 1966, **1**, 1-19.

Bijou, S. W., & Baer, D. M. *Child development: Readings in experimental analysis*. New York: Appleton, 1967.

Bijou, S. W., & Orlando, R. Rapid development of multiple-schedule performances with retarded children. *Journal of the Experimental Analysis of Behavior*, 1961, **4**, 7-16.

Binder, A., McConnell, D., & Sjoholm, N. A. Verbal conditioning as a function of experimenter characteristics. *Journal of Abnormal and Social Psychology*, 1957, **55**, 309-314.

Birnbrauer, J. S., Bijou, S. W., Wolf, M. M., & Kidder, J. D. Programmed instruction in the classroom. In L. Ullmann & L. Krasner (Eds.), *Case studies in behavior modification*. New York: Holt, 1965. Pp. 358-363.

Birnbrauer, J. S., & Lawler, J. Token reinforcement for learning. *Mental Retardation*, 1964, **2**, 275-279.

Blackman, L. S., & Smith, M. P. *The development and evaluation of a curriculum for educable mental retardates utilizing self-instructor devices or teaching machines*. Bordentown, New Jersey: Johnstone Learning and Research Center, 1964.

Blank, J. P. The effect of magnitude of incentive on acquisition and extinction in mentally retarded children. *Dissertation Abstracts*, 1958, **19**, 1443.

Blank, M., & Altman, L. Effect of stimulus modality and task complexity on discrimination and reversal learning in preschool children. *Journal of Experimental Child Psychology*, 1968, **6**, 598-606.

Blehert, S. L. Pattern discrimination learning with rhesus monkeys. Technical Report No. 77, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965.

Bliss, E. L. (Ed.) *Roots of behavior: Genetics, instinct, and socialization in animal behavior*. New York: Harper (Hoeber), 1962.

Blue, C. M. Performance of normal and retarded subjects on a modified paired-associate task. *American Journal of Mental Deficiency*, 1963, **68**, 228-234.

Bousfield, W. A. The occurrences of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, 1953, **49**, 229-240.

Bousfield, W. A., Esterson, J., & Whitmarsh, G. A. A study of developmental changes in conceptual and perceptual associative clustering. *Journal of Genetic Psychology*, 1958, **92**, 95-102.

Bovet, D., Bovet-Nitti, F., & Oliverio, A. Genetic aspects of learning and memory in mice. *Science*, 1969, **163**, 139-149.

Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*, 1967, **1**, 229-325.

Bower, G. H., & Grusec, T. Effect of prior Pavlovian discrimination training upon learning an operant discrimination. *Journal of the Experimental Analysis of Behavior*, 1964, **7**, 401-404.

Bower, G. H., Starr, R., & Lazarovitz, L. Amount of response-produced change in CS and avoidance learning. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 13-17.

Brackbill, Y. Extinction of the smiling response in infants as a function of reinforcement schedule. *Child Development*, 1958, **29**, 115-124.

Broadbent, D. E. *Perception and communication*. Oxford: Pergamon Press, 1958.

Brown, J. S., Bilodeau, E. A., & Baron, M. R. Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual-spatial dimension. *Journal of Experimental Psychology*, 1951, **41**, 52-61.

Brown, Z., and Smith, J. Tactual-kinetic discrimination shifts in retarded children. *Journal of Experimental Child Psychology*, 1967, **5**, 455-462.

Bryant, P. E. The effects of verbal labelling on recall and recognition in severely subnormal and normal children. *Journal of Mental Deficiency Research*, 1965, **9**, 229-236. (a)

Bryant, P. E. The effects of verbal labelling on recognition of pictures and names in severely subnormal and normal children. *Journal of Mental Deficiency Research*, 1965, **9**, 237-244. (b)

Bush, R. R., & Mosteller, F. *Stochastic models for learning*. New York: Wiley, 1955.

Caldwell, E. C., & Hall, V. C. The influence of concept training on letter discriminations. *Child Development*, 1969, **40**, 63-71.

Calfee, R. C. Short-term retention in normal and retarded children as a function of memory load and list structure. Technical Report No. 75, University of Wisconsin, 1969.

Campione, J., Hyman, L., & Zeaman, D. Dimensional shifts and reversals in retardate discrimination learning. *Journal of Experimental Child Psychology*, 1965, **2**, 255-263.

Cantor, G. N. Hull-Spence behavior theory and mental deficiency. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 92-133.

Capaldi, E. J. A sequential hypothesis of instrumental learning. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation*, 1967, **1**, 67-156.

Caron, A. J. Conceptual transfer in preverbal children as a consequence of dimensional training. *Journal of Experimental Child Psychology*, 1968, **6**, 522-542.

Church, R. M., & Campbell, B. A. (Eds.) *Punishment and avoidance behavior*. New York: Appleton, 1969.

Cofer, C. N., & Musgrave, B. (Eds.). *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963.

Cole, M., Frankel, F., and Sharp, D. The development of free recall learning in children. *Developmental Psychology*, 1970, **2**, in press.

Cole, M., & Maltzman, I. (Eds.) *A handbook of contemporary Soviet psychology*. New York: Basic Books, 1969.

Cook, S. W., & Harris, R. E. The verbal conditioning of the galvanic skin reflex. *Journal of Experimental Psychology*, 1937, **21**, 202-210.

Craig, G. J., & Myers, J. L. A developmental study of sequential two-choice decision making. *Child Development*, 1963, **34**, 483-493.

Cromwell, R. L. A social learning approach to mental retardation. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 41-91.

Cromwell, R. L., Palk, B. E., & Foshee, J. G. Studies in activity level: V. The relationships among eyelid conditioning, intelligence activity level, and age. *American Journal of Mental Deficiency*, 1961, **65**, 744-748.

Cruickshank, W. M., & Blake, K. A. A comparative study of the performance of mentally handicapped and intellectually normal boys on selected tasks involving learning and transfer. Syracuse, New York: University of Syracuse Press, 1957. Cooperative Research Project SAE-6414.

Cunningham, M. A. The effects of training procedures on visual discrimination learning and transfer in retarded and normal children. *Dissertation Abstracts*, 1965, **26**, 1180-1181.

D'Amato, M. R., & Jagoda, H. Analysis of the role of overlearning in discrimination reversal. *Journal of Experimental Psychology*, 1961, **61**, 45-50.

Davis, R. E., Bright, P. I., & Agranoff, B. W. Effect of ECS and puromycin on memory in fish. *Journal of Comparative and Physiological Psychology*, 1965, **60**, 162-166.

Denny, M. R. Learning through stimulus satiation. *Journal of Experimental Psychology*, 1957, **54**, 62-64.

Denny, M. R. Research in learning performance. In H. Stevens & R. Heber (Eds.), *Mental retardation*. Chicago, Illinois: University Press, 1964. Pp. 100-142.

Denny, M. R. A theoretical analysis and its application to training the mentally retarded. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*, 1966, **2**, 1-27.

Dickerson, D. J. Performance of preschool children on four discrimination shifts. *Journal of Experimental Child Psychology*, 1967, **5**, 228-236.

Dickerson, D. J., Girardeau, F. L., & Spradlin, J. E. Verbal pretraining and discrimination learning by retardates. *American Journal of Mental Deficiency*, 1964, **68**, 476-484.

Doubros, S. G. Sensory preconditioning with the mentally retarded. *American Journal of Mental Deficiency*, 1965, **70**, 282-286.

Egger, M. D., & Miller, N. E. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology*, 1962, **64**, 97-104.

Eimas, P. D. Optional shift behavior in children as a function of overtraining, irrelevant stimuli, and age. *Journal of Experimental Child Psychology*, 1967, **5**, 332-340.

Eisman, B. S. Paired associate learning, generalization, and retention as a function of intelligence. *American Journal of Mental Deficiency*, 1958, **63**, 481-489.

Ellis, N. R. The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 134-158.

Ellis, N. R. (Ed.) *Handbook of mental deficiency*. New York: McGraw-Hill, 1963.

Ellis, N. R. (Ed.) *International review of research in mental retardation*, 1966, **1**. (a)

Ellis, N. R. (Ed.) *International review of research in mental retardation*, 1966, **2**. (b)

Ellis, N. R. (Ed.) *International review of research in mental retardation*, 1968, **3**.

Ellis, N. R. A behavioral research strategy in mental retardation: defense and critique. *American Journal of Mental Deficiency*, 1969, **73**, 557-566.

Ellis, N. R. Memory processes in retardates and normals: Theoretical and empirical considerations. In N. R. Ellis (Ed.), *International review of research in mental retardation*, 1970, **4**.

Ellis, N. R. (Ed.) *International review of research in mental retardation*, 1970, **4**.

Ellis, N. R., Barnett, C. D., & Pryer, M. W. Operant behavior in mental defectives: exploratory studies. *Journal of the Experimental Analysis of Behavior*, 1960, **3**, 63-69.

Ellis, N. R., & Hope, R. Memory processes and the serial position curve. *Journal of Experimental Psychology*, 1968, **77**, 613-619.

Ellis, N. R., Pryer, M. W., & Barnett, C. D. Motor learning and retention in normals and defectives. *Perceptual and Motor Skills*, 1960, **10**, 83-91.

Ellis, N. R., Pryer, M. W., Distefano, M., & Pryer, R. S. Learning in mentally defective, normal, and superior subjects. *American Journal of Mental Deficiency*, 1959, **64**, 725-734.

Ellis, N. R., & Sloan, W. Rotary pursuit performance as a function of mental age. *Perceptual and Motor Skills*, 1957, **7**, 267-270.

Ellis, N. R., & Sloan, W. Oddity learning as a function of mental age. *Journal of Comparative and Physiological Psychology*, 1959, **56**, 228-230.

Ellson, D. G., Harris, P., & Barber, L. A field test of programmed and directed tutoring. *Reading Research Quarterly*, 1968, **3**, 307-367.

Estes, W. K. Spontaneous recovery from extinction in maze-bright and maze-dull rats. *Journal of Comparative and Physiological Psychology*, 1942, **34**, 349-351.

Estes, W. K. Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *Journal of Experimental Psychology*, 1948, **38**, 173-177.

Estes, W. K. Toward a statistical theory of learning. *Psychological Review*, 1950, **57**, 94-107.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, **62**, 145-154.

Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A Study of a Science*. 1959, **2**, 380-491.

Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 89-128.

Estes, W. K. Reinforcement in human learning. In J. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969. Pp. 63-94.

Estes, W. K. Outline of a theory of punishment. Technical Report No. 123, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1967. In R. M. Church & B. A. Campbell (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969. Pp. 57-82.

Estes, W. K. New perspectives on some old issues in association theory. In W. K. Honig & N. J. Mackintosh (Eds.), *Fundamental issues in associative learning*. Halifax: Dalhousie University Press, 1969. Pp. 162-189.

Estes, W. K., Koch, S., MacCorquodale, K., Meehl, P. E., Mueller, C. G., Jr., Schoenfeld, W. N., & Verplanck, W. S. *Modern learning theory*. New York: Appleton, 1954.

Estes, W. K., & Lauer, D. W. Conditions of invariance and modifiability in simple reversal learning. *Journal of Comparative and Physiological Psychology*, 1957, **50**, 199-206.

Estes, W. K., & Schoeffler, M. S. Analysis of variables influencing alternation after forced trials. *Journal of Comparative and Physiological Psychology*, 1955, **48**, 357-362.

Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *Journal of Experimental Psychology*, 1941, **29**, 390-400.

Fernald, G. M. *Remedial techniques in basic school subjects*. New York: McGraw-Hill, 1943.

Ferster, C. B., & Skinner, B. F. *Schedules of reinforcement*. New York: Appleton, 1957.

Fields, P. E. Multiple discriminations learning by white rats. *Journal of Comparative and Physiological Psychology*, 1954, **47**, 472-476.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, 1966, **37**, 283-299.

Franks, V., & Franks, C. M. Conditionability in defectives and in normals as related to intelligence and organic deficit: The application of a learning theory model to the study of the learning process in the mental defective. In *Proceedings of Conference on the Scientific Study of Mental Deficiency, London, 1960*, 1962, **2**, 577-583.

Fuller, J. L. Hereditary differences in trainability of purebred dogs. *Journal of Genetic Psychology*, 1955, **89**, 229-238.

Fuller, J. L., & Thompson, W. R. *Behavior genetics*. New York: Wiley, 1960.

Fuller, P. R. Operant conditioning of a vegetative human organism. *American Journal of Psychology*, 1949, **62**, 587-589.

Furth, H. G., & Youniss, J. Effect of overtraining on three discrimination shifts in children. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 290-293.

Gagné, R. M. *Learning and individual differences*. Columbia, Ohio: Merrill Books, 1967.

Gagné, R. M. Contributions of learning to human development. *Psychological Review*, 1968, **75**, 177-191.

Garcia, J., & Koelling, R. A. Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 1966, **4**, 123-124.

Garcia, J., McGowan, B. K., Ervin, F. R., & Koelling, R. A. Cues: Their relative effectiveness as a function of the reinforcer. *Science*, 1968, **160**, 794-795.

Gardner, W. I. Effects of failure on intellectually retarded and normal boys. *American Journal of Mental Deficiency*, 1966, **70**, 899-902.

Gewirtz, J. L. Mechanisms of social learning: Some roles of stimulation and behavior in early human development. In D. A. Goslin (Ed.), *Handbook of socialization theory and research*. New York: Rand McNally, 1969. Pp. 57-212.

Gibson, E. J. Learning to read. *Science*, 1965, **148**, 1066-1072.

Girardeau, F. L., & Ellis, N. R. Rote verbal learning by normal and mentally retarded children. *American Journal of Mental Deficiency*, 1964, **68**, 525-532.

Girardeau, F. L., & Spradlin, J. E. Token rewards in a cottage program. *Mental Retardation*, 1964, **2**, 345-351.

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), *Learning and individual differences*. Columbia, Ohio: Merrill Books, 1967. Pp. 1-22.

Gormezano, I., & Moore, J. W. Effects of instrumental set and UCS intensity on the latency, percentage, and form of the eyelid response. *Journal of Experimental Psychology*, 1962, **53**, 245-250.

Goulet, L. R. Verbal learning and memory research with retardates: An attempt to assess developmental trends. In N. R. Ellis (Ed.), *International review of research in mental retardation*, 1968, **3**, 97-134.

Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *Journal of Experimental Psychology*, 1948, **38**, 690-696.

Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *Journal of Experimental Psychology*, 1949, **39**, 35-40.

Gray, S. W., & Klaus, R. A. An experimental preschool program for culturally deprived children. Unpublished paper given at American Association for Advancement of Science, Montreal, 1964.

Greenberg, R., & Underwood, B. J. Retention as a function of stage of practice. *Journal of Experimental Psychology*, 1950, **40**, 452-457.

Greeno, J. G. *Elementary theoretical psychology*. Reading, Massachusetts: Addison-Wesley, 1969.

Griffith, B. C., Spitz, H. H., & Lipman, R. S. Verbal mediation and concept formation in retarded and normal subjects. *Journal of Experimental Psychology*, 1959, **58**, 247-251.

Grings, W. W., Lockhart, R. S., & Dameron, L. E. Conditioning autonomic responses of mentally subnormal individuals. *Psychological Monographs*, 1962, **76**, No. 39, 1-35.

Grossman, S. P. *A textbook of physiological psychology*. New York: Wiley, 1967.

Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.

Guthrie, E. R. *The psychology of learning*. New York: Harper, 1935.

Guthrie, E. R. *The psychology of learning*. (Rev. ed.) New York: Harper, 1952.

Hall, J. F. *The psychology of learning*. Philadelphia, Pennsylvania: Lippincott, 1966.

Hamilton, J. Learning of a generalized response class in mentally retarded individuals. *American Journal of Mental Deficiency*, 1966, **71**, 100-108.

Harlow, H. F. The formation of learning sets. *Psychological Review*, 1949, **56**, 51-65.

Harlow, H. F. The development of learning in the rhesus monkey. *American Scientist*, 1959, **47**, 459-479. (a)

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2, New York: McGraw-Hill, 1959. Pp. 492-537. (b)

Harris, J. D. Forward conditioning, backward conditioning, pseudoconditioning, and adaptation to the conditioned stimulus. *Journal of Experimental Psychology*, 1941, **28**, 491-502.

Harvey, O. J. (Ed.) *Experience, structure, and adaptability*. New York: Springer, 1966.

Hayes, K. J. The backward curve: A method for the study of learning. *Psychological Review*, 1953, **60**, 269-275.

Haywood, C. (Ed.) *Psychometric intelligence*. New York: Appleton, in press.

Haywood, H. C., & Tapp, J. T. Experience and the development of adaptive behavior. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*, 1966, **1**, 109-151.

Heath, R. G. (Ed.). *The role of pleasure in behavior*. New York: Hoeber, 1964.

Hebb, D. O. *The organization of behavior: A neuropsychological theory*. New York: Wiley, 1949.

Heber, R. F. Expectancy and expectancy changes in normal and mentally retarded boys. Ann Arbor, Michigan: University Microfilms, 1957.

Helson, H., & Bevan, W. (Eds.) *Contemporary approaches to psychology*. Princeton, New Jersey: Van Nostrand, 1967.

Heron, W. T. Inheritance of maze-learning ability in rats. *Journal of Comparative Psychology*, 1935, **19**, 77-89.

Heron, W. T. The inheritance of brightness and dullness in maze learning ability in the rat. *Journal of Genetic Psychology*, 1941, **59**, 41-49.

Heron, W. T., & Skinner, B. F. The rate of extinction in maze-bright and maze-dull rats. *Psychological Record*, 1940, **4**, 11-18.

Hilgard, E. R. An algebraic analysis of conditioned discrimination in man. *Psychological Review*, 1938, **45**, 472-496.

Hilgard, E. R., & Atkinson, R. C. *Introduction to psychology*. (3rd ed.) New York: Harcourt, 1967.

Hilgard, E. R., & Bower, G. H. *Theories of learning* (3rd ed.). New York: Appleton, 1966.

Hill, S. D. The performance of young children on three discrimination-learning tasks. *Child Development*, 1965, **36**, 423-435.

Hirsch, J. Individual differences in behavior and their genetic bases. In E. L. Bliss (Ed.), *Roots of behavior: Genetic, instinct, and socialization in animal behavior*. New York: Hoeber, 1962. Pp. 3-23.

Hirsch, J., & Erlenmeyer-Kimling, L. Behavior genetics. In H. Helson & W. Bevan (Eds.), *Contemporary approaches to psychology*. Princeton, New Jersey: Van Nostrand, 1967. Pp. 91-137.

Honig, W. K., & Mackintosh, N. J. (Eds.) *Fundamental issues in associative learning*. Halifax, Nova Scotia: Dalhousie University Press, 1969.

House, B. J., Smith, M., & Zeaman, D. Verbal learning and retention as a function of number of lists in retardates. *American Journal of Mental Deficiency*, 1964, **69**, 239-243.

House, B. J., & Zeaman, D. Reward and nonreward in the discrimination learning of imbeciles. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 614-618.

House, B. J., & Zeaman, D. Position discrimination and reversals in low-grade retardates. *Journal of Comparative and Physiological Psychology*, 1959, **52**, 564-565.

House, B. J., & Zeaman, D. Visual discrimination learning and intelligence in defectives of low mental age. *American Journal of Mental Deficiency*, 1960, **65**, 51-58. (a)

House, B. J., & Zeaman, D. Transfer of a discrimination from objects to patterns. *Journal of Experimental Psychology*, 1960, **59**, 298-302. (b)

House, B. J., & Zeaman, D. Reversal and nonreversal shifts in discrimination learning in retardates. *Journal of Experimental Psychology*, 1962, **63**, 444-451.

Hull, C. L. *Principles of behavior*. New York: Appleton, 1943.

Hull, C. L. The place of innate individual and species difference in a natural-science theory of behavior. *Psychological Review*, 1945, **52**, 55-60.

Hull, C. L. *Essentials of behavior*. New Haven, Connecticut: Yale University Press, 1951.

Irwin, O. C. Amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life. *Genetic Psychology Monographs*, 1930, **8**, 1-92.

Jacobs, A. Performance of children in a discrimination problem as a function of symbolic guidance, delay of reward, and mental ability. Unpublished doctoral dissertation, State University of Iowa, 1950.

Jarvik, M. E. Discrimination of colored food and food signs by primates. *Journal of Comparative and Physiological Psychology*, 1953, **46**, 390-392.

Jeffrey, W. E., & Cohen, L. B. Effect of spatial separation of stimulus, response, and reinforcement on selective learning in children. *Journal of Experimental Psychology*, 1964, **67**, 577-580.

Jenkins, H. M., & Sainsbury, R. S. The development of stimulus control through differential reinforcement. Technical Report No. 27, Department Psychology, McMasters University, 1968. In W. K. Honig & N. J. Mackintosh (Eds.), *Fundamental issues in associative learning*. Halifax, Nova Scotia: Dalhousie University Press, 1969. Pp. 123-161.

Jensen, A. R. Learning abilities in retarded, average, and gifted children. *Merrill-Palmer Quarterly*, 1963, **9**, 123-150.

Jensen, A. R. Rote learning in retarded adults and normal children. *American Journal of Mental Deficiency*, 1965, **69**, 828-834.

Jensen, A. R., & Rohwer, W. D., Jr. The effect of verbal mediation on the learning and retention of paired associates by retarded adults. *American Journal of Mental Deficiency*, 1963, **68**, 80-84. (a)

Jensen, A. R., & Rohwer, W. D., Jr. Verbal mediation in paired-associate and serial learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, **1**, 346-352. (b)

John, E. R., Chesler, P., Bartlett, F., & Victor, I. Observation learning in cats. *Science*, 1968, **159**, 1489-1491.

Johnson, B. M. Reward schedules and instrumental conditioning in normal and retarded children. *Child Development*, 1966, **37**, 633-644.

Jones, M. R. (Ed.) *Nebraska symposium on motivation, 1963*. Lincoln, Nebraska: Nebraska University Press, 1963.

Jones, R. W., & Ellis, N. R. Inhibitory potential in rotary pursuit acquisition by normal and defective subjects. *Journal of Experimental Psychology*, 1962, **63**, 534-537.

Kamin, L. J. Temporal and intensity characteristics of the conditioned stimulus. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium*. New York: Appleton, 1965. Pp. 118-147.

Kamin, L. J. 'Attention-like' processes in classical conditioning. Technical Report No. 5, Department of Psychology, McMaster University, 1967.

Kaufman, M. E., & Peterson, W. M. Acquisition of a learning set by normal and mentally retarded children. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 619-621.

Kaufman, M. E., & Peterson, W. M. Acquisition of a conditional discrimination learning-set by normal and mentally retarded children. *American Journal of Mental Deficiency*, 1965, **69**, 865-890.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. *Psychological Review*, 1962, **69**, 1-16.

Kendler, H. H., Kendler, T. S., & Marken, R. S. Developmental analysis of reversal and half-reversal shifts. *Developmental Psychology*, 1969, **1**, 318-326.

Kendler, T. S. Development of mediating responses in children. In J. C. Wright & J. Kagan (Eds.), Basic cognitive processes in children. *Monographs of Social Research and Child Development*, 1962, **28**, No. 2, Serial 86, 33-52.

Kendler, T. S., & Kendler, H. H. Reversal and nonreversal shifts in kindergarten children. *Journal of Experimental Psychology*, 1959, **58**, 56-60.

Keppel, G. Verbal learning in children. *Psychological Bulletin*, 1964, **61**, 63-80.

Kimble, G. A. *Hilgard & Marquis' conditioning and learning*. (2nd ed.) New York: Appleton, 1961.

Kingsley, P. R., & Hagen, J. W. Induced versus spontaneous rehearsal in short-term memory in nursery school children. *Developmental Psychology*, 1969, **1**, 40-46.

Koch, S. (Ed.) *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959.

Kohler, W. *Gestalt psychology*. New York: Horace Liveright, 1929.

Kohler, W., & Wallach, H. Figural after-effects: An investigation of visual processes. *Proceedings of the American Philosophical Society*, 1944, **88**, 269-357.

Kounin, J. Experimental studies of rigidity. II. The explanatory power of the concept of rigidity as applied to feeblemindedness. *Character and Personality*, 1941, **9**, 273-282.

Krasner, L. Studies of the conditioning of verbal behavior. *Psychological Bulletin*, 1958, **55**, 148-170.

Krasner, L., & Ullmann, L. P. (Eds.) *Research in behavior modification: New developments and implications.* New York: Holt, 1965.

Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *Journal of Experimental Psychology*, 1946, **36**, 471-490.

Lashley, K. S. The mechanism of vision. I. A method for rapid analysis of pattern-vision in the rat. *Journal of Genetic Psychology*, 1930, **37**, 453-460.

Laurence, M. W. Age differences in performance and subjective organizations in the free-recall learning of pictorial material. *Canadian Journal of Psychology*, 1966, **20**, 388-399.

Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, 1949, **39**, 770-784.

Lawrence, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 1950, **40**, 175-188.

Lawrence, D. H. The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, 1952, **45**, 511-516.

Lawrence, D. H. The nature of a stimulus: some relationships between learning and perception. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 5. New York: McGraw-Hill, 1963. Pp. 179-212.

Lawson, R., & Watson, L. J. Transfer of training from a machine program by mentally retarded children. Mimeographed report. Columbus, Ohio: Ohio State University, College of Education, 1964.

Lenneberg, E. H. *Biological foundations of language.* New York: Wiley, 1967.

Leontiev, A. N. The nature and formation of human psychic properties. In B. Simon (Ed.), *Psychology in the Soviet Union.* London: Routledge & Kegan Paul, 1957. Pp. 226-232.

Levine, M. Cue neutralization: The effects of random reinforcement upon discrimination learning. *Journal of Experimental Psychology*, 1962, **63**, 438-443.

Levine, M. Mediating processes in humans at the outset of discrimination learning. *Journal of Experimental Psychology*, 1963, **70**, 254-276.

Lewin, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

Lindsley, O. R. Operant conditioning methods applied to research in chronic schizophrenia. *Psychiatric Research Reports*, 1956, **5**, 118-139.

Lipman, R. S. Learning: Verbal, perceptual-motor, and classical conditioning. In N. R. Ellis (Ed.), *Handbook of mental deficiency.* New York: McGraw-Hill, 1963. Pp. 391-423.

Lipsitt, L. P. Learning in the first year of life. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in Child Development and Behavior.* 1963, **1**, 147-195.

Lipsitt, L. P., & Spiker, C. C. (Eds.) *Advances in Child Development and Behavior*, 1963, **1**.

Lipsitt, L. P., & Spiker, C. C. (Eds.) *Advances in Child Development and Behavior*, 1965, **2**.

Lipsitt, L. P., & Spiker, C. C. (Eds.). *Advances in Child Development and Behavior*, 1967, **3**.

Locke, B. J. Verbal conditioning with the retarded: Reinforcer, sex of subject, and stimulus pacing. *American Journal of Mental Deficiency*, 1968, **73**, 616-620. (a)

Locke, B. J. Verbal conditioning with retarded subjects: Establishment or reinstatement of effective reinforcing consequences. *American Journal of Mental Deficiency*, 1968, **73**, 621-626. (b)

Lovaas, O. J., Freitas, L., Guilani, B., Nelson, K., & Whalen, C. Building of social and preschool behavior in schizophrenic children through nonverbal imitation training. *Behavior Research and Therapy*, 1967, **5**, 171-181.

Lovaas, O. J., Schaeffer, B., & Simmons, J. Q. Building social behavior in autistic children by use of electric shock. *Journal of Experimental Research on Personality*, 1965, **1**, 99-109.

Lovejoy, E. An attention theory of discrimination learning. *Journal of Mathematical Psychology*, 1965, **2**, 342-362.

Lovejoy, E. Analysis of the overlearning reversal effect. *Psychological Review*, 1966, **73**, 87-103.

Lumsdaine, A. A., & Glaser, R. *Teaching machines and programmed learning*. Washington, D. C.: National Education Association, 1960.

Luria, A. R. *The mentally retarded child: Essays based on a study of the peculiarities of higher nervous functions*. New York: Macmillan, 1963.

MacCorquodale, K., Meehl, P., & Tolman, E. C. In W. K. Estes *et al*. (Eds.), *Modern learning theory*, New York: Appleton, 1954. Pp. 177-266.

Mackintosh, N. J. The effect of attention on the slope of generalization gradients. *British Journal of Psychology*, 1965, **56**, 87-93.

Malpass, L. F., Hardy, M. W., Gilmore, A. S., & Williams, C. F. Automated instruction for retarded children. *American Journal of Mental Deficiency*, 1964, **69**, 405-412.

Marquis, D. P. Can conditioned responses be established in the newborn infant? *Journal of Genetic Psychology*, 1931, **39**, 479-492.

McClearn, G. E., & Meredith, W. Behavioral genetics. *Annual Review of Psychology*, 1966, **18**, 515-550.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton, 1953.

McCrary, J. W., & Hunter, W. S. Serial position curves in verbal learning. *Science*, 1953, **117**, 131-134.

McGeoch, J. A., & Irion, A. L. *The psychology of human learning*. New York: Longmans, Green, 1952.

Melton, A. W. (Ed.) *Categories of human learning*. New York: Academic Press, 1964.

Melton, A. W. Individual differences and theoretical process variables: General comments on the Conference. In R. M. Gagné (Ed.), *Learning and individual differences*. Columbus, Ohio: Merrill Books, 1967. Pp. 238-252.

Metz, J. R. Conditioning social and intellectual skills in autistic children. *California Conference on Reinforcement Theory in Psychological Treatment*, 1966. Pp. 40-49.

Meyer, D. R., & Miles, R. C. Intralist-interlist relations in verbal learning. *Journal of Experimental Psychology*, 1953, **45**, 109-115.

Milgram, N. A. Retention of mediation set in paired-associate learning of normal children and retardates. *Journal of Experimental Child Psychology*, 1967, **5**, 341-349.

Milgram, N. A., & Furth, H. G. Position reversal vs. dimension reversal in normal and retarded children. *Child Development*, 1964, **35**, 701-708.

Milgram, N. A., & Noce, J. S. Relevant and irrelevant verbalization in discrimination and reversal learning by normal and retarded children. *Journal of Educational Psychology*, 1968, **59**, 169-175.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.

Miller, J. The effect of facilitatory and inhibitory attitudes on eyelid conditioning. Unpublished doctoral dissertation, Yale University, 1939.

Miller, M. B. Locus of control, learning climate, and climate shift in serial learning with retardates. *Dissertation Abstracts*, 1962, **22**, 2887.

Miller, N. E. Liberalization of Basic S-R Concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A Study of a Science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 196-292.

Miller, N. E. Some reflections on the law of effect produce a new alternative to drive reduction. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, Lincoln, Nebraska: Nebraska University Press, 1963. Pp. 65-112.

Milner, B. Amnesia following operation on the temporal lobes. In C. W. M. Whitty & O. L. Zangwell (Eds.), *Amnesia*. Washington, D. C., and London: Butterworth, 1966. Pp. 109-133.

Moely, B. E., Olson, F. A., Hawles, T. G., & Flavell, J. H. Production deficiency in young children's clustered recall. *Developmental Psychology*, 1969, **1**, 26-34.

Montessori, M. *The Montessori method*. New York: Frederick A. Stokes, 1912.

Moore, O. K. Autotelic responsive environments and exceptional children. In O. J. Harvey (Ed.), *Experience, structure, and adaptability*. New York: Springer, 1966. Pp. 169-216.

Muenzinger, K. F., & Fletcher, F. M. Motivation in learning. VII. The effect of an enforced delay at the point of choice in the visual discrimination habit. *Journal of Comparative Psychology*, 1937, **23**, 383-392.

Murphy, J. V., & Miller, R. E. The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. *Journal of Comparative and Physiological Psychology*, 1955, **48**, 221-224.

Neimark, E. D., & Estes, W. K. (Eds.) *Stimulus sampling theory*. San Francisco, California: Holden-Day, 1967.

Noble, C. E., & Fuchs, J. E. Serial errors in human learning: a test of the McCrary-Hunter hypothesis. *Science*, 1959, **129**, 570-571.

Oatley, K. G., Bryant, P. E., & Tinson, C. Non-reinforcement and the emission of alternate responses by severely subnormal children. *Journal of Mental Deficiency Research*, 1965, **9**, 191-200.

Olds, J. Differential effect of drive and drugs on self-stimulation at different brain sites. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin, Texas: Texas University Press, 1961. Pp. 350-366.

Olton, D. S., & Isaacson, R. L. Importance of spatial location in active avoidance tasks. *Journal of Comparative and Physiological Psychology*, 1968, **65**, 535-539.

Orlando, R. Shaping multiple schedule performance in retardates: Establishment of baselines by systematic and special procedures. *Journal of Experimental Child Psychology*, 1965, **2**, 135-153.

Orlando, R., & Bijou, S. W. Single and multiple schedules of reinforcement in developmentally retarded children. *Journal of the Experimental Analysis of Behavior*, 1960, **3**, 339-348.

Orlando, R., & Bijou, S. W. Rapid acquisition of multiple-schedule discrimination performance in a single response free-operant situation with retardates. In S. W. Bijou & B. M. Baer, *Child Development: Readings in experimental analysis*. New York: Appleton, 1967. Pp. 81-92.

Paraskevopoulos, I. Symmetry, recall, and preference in relation to chronological age. *Journal of Experimental Child Psychology*, 1968, **6**, 254-264. (a)

Paraskevopoulos, I. Developmental stages for decoding symmetry in retarded and gifted children. *American Journal of Mental Deficiency*, 1968, **73**, 447-454. (b)

Passey, G. E. The influence of intensity of unconditioned stimulus on acquisition of a conditioned response. *Journal of Experimental Psychology*, 1948, **38**, 420-428.

Pavlov, I. P. *Conditioned Reflexes* (Translated by G. V. Anrep). London and New York: Oxford University Press, 1927.

Pennington, D. F., & Thompson, R. SR: Discrimination learning in rats as a function of highly distributed trials. *Journal of Experimental Psychology*, 1958, **56**, 94-95.

Peterson, L. R. Immediate memory: data and theory. In C. N. Cofer & B. Musgrave (Eds.), *Verbal behavior and learning: problems and processes*. New York: McGraw-Hill, 1963. Pp. 336-353.

Phillips, J. L., Shiffrin, R. M., & Atkinson, R. C. The effects of list length on short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 303-311.

Piaget, J. *The language and thought of the child*. New York: Harcourt, 1926.

Piaget, J. *The origin of intelligence in the child*. New York: International University Press, 1952.

Piaget, J. *On the development of memory and identity*. Heinz Werner Lectures, Clark University. Vol. 2. Barre, Massachusetts: Barre Publ., 1967.

Postman, L. Short-term memory and incidental learning. In A. W. Melton (Ed.) *Categories of human learning*. New York: Academic Press, 1964. Pp. 145-201.

President's Panel on Mental Retardation. A Proposed Program for National Action to Combat Mental Retardation. Washington, D. C.: U.S. Government Printing Office, 1962. Pp. 201+vii.

Pressey, S. L. A simple apparatus which gives tests and scores—and teaches. *School & Society*, 1926, **23**, 373-376.

Price, J. E. Automated teaching program with mentally retarded students. *American Journal of Mental Deficiency*, 1963, **68**, 69-72.

Prokasy, W. F. (Ed.) *Classical conditioning: A symposium*. New York: Appleton, 1965.

Ramsey, O. E. The effects of four instructional procedures on free operant discrimination reversal in retardates. *Dissertation Abstracts*, 1963, **24**, 406.

Rescorla, R. A. Predictability and number of pairings in Pavlovian fear conditioning. *Psychonomic Science*, 1966, **4**, 383-384.

Rescorla, R. A. Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 1968, **66**, 1-5.

Restle, F. A theory of discrimination learning. *Psychological Review*, 1955, **62**, 11-19.

Ringelheim, D. The effect of failure on verbal expectancies in male defectives. *Abstract of Peabody Studies in Mental Retardation*, 1960, **1**, No. 14.

Ross, L. E. Classical conditioning and discrimination learning research with the mentally retarded. In N. R. Ellis (Ed.) *International Review of Research in Mental Retardation*, 1966, **1**, 21-54.

Ross, L. E., Koski, C. H., & Yager, J. Classical eyelid conditioning of the severely retarded: Partial reinforcement effects. *Psychonomic Science*, 1964, **1**, 253-254.

Rotter, J. B. *Social learning and clinical psychology*. Englewood Cliffs, New Jersey: Prentice Hall, 1954.

Rudel, R. G. The absolute response in tests of generalization in normal and retarded children. *American Journal of Psychology*, 1959, **72**, 401-408.

Sanders, B., Ross, L. E., & Heal, L. W. Reversal and nonreversal shift learning in normal children and retardates of comparable mental age. *Journal of Experimental Psychology*, 1965, **69**, 84-88.

Scott, K. G. Intelligence and Learning. In C. Haywood (Ed.), *Psychometric intelligence*. New York: Appleton, in press.

Sheer, D. E. (Ed.) *Electrical stimulation of the brain*. Austin, Texas: University of Texas Press, 1961.

Shepp, B. E. Some cue properties of incentives: Discrimination of distinct rewards by retardates. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 1078-1080.

Shepp, B. E. Some cue properties of rewards in simultaneous object-discriminations of retardates. *Child Development*, 1964, **35**, 587-592.

Shepp, B. E., & Eimas, P. D. Intradimensional and extradimensional shifts in the rat. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 357-361.

Shepp, B. E., & Turrisi, F. D. Learning and transfer of mediating responses in discrimination learning. In N. R. Ellis (Ed.) *International Review of Research in Mental Retardation*, 1966, **2**, 85-121.

Shepp, B. E., & Zeaman, D. Discrimination learning of size and brightness by retardates. *Journal of Comparative and Physiological Psychology*, 1966, **62**, 55-59.

Shif, Z. I. Development of children in schools for the mentally retarded. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology*. New York: Basic Books, 1969. Pp. 326-353.

Sidman, M., & Stoddard, L. T. Programming perception and learning for retarded children. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*, 1966, **2**, 151-208.

Siegel, P. S., & Foshee, J. G. Molar variability in the mental defective. *Journal of Abnormal and Social Psychology*, 1960, **61**, 141-143.

Simmons, M. W. Operant discrimination learning in human infants. *Child Development*, 1964, **35**, 737-748.

Simon, B. *Psychology in the Soviet Union*. London: Routledge & Kegan Paul, 1957.

Skinner, B. F. *Behavior of organisms*. New York: Appleton, 1938.

Skinner, B. F. Are theories of learning necessary? *Psychological Review*, 1950, **57**, 193-216.

Skinner, B. F. The science of learning and the art of teaching. *Harvard Educational Review*, 1954, **24**, 86-97.

Skinner, B. F. Teaching machines. *Science*, 1958, **128**, 969-977.

Slavina, L. S. Specific features of the intellectual work of unsuccessful pupils. In B. Simon (Ed.), *Psychology in the Soviet Union*. London: Routledge & Kegan Paul, 1957. Pp. 205-212.

Smith, M. P., Means, J. R., & Fishkin, S. Effects of naming upon learning and transfer of learning in the mentally retarded. *American Journal of Mental Deficiency*, 1968, **72**, 637-647.

Solomon, R. L., & Turner, L. H. Discriminative classical conditioning in dogs paralyzed by curare can later control discriminative avoidance responses in the normal state. *Psychological Review*, 1962, **69**, 202-219.

Spearman, C. *The abilities of man*. New York: Macmillan, 1927.

Spence, J. T., & Bower, G. H. (Eds.) *The Psychology of Learning and Motivation*, 1968, **2**.

Spence, K. W. The nature of discrimination learning in animals. *Psychological Review*, 1936, **43**, 427-449.

Spence, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 690-729.

Spence, K. W. *Behavior theory and conditioning*. New Haven, Connecticut: Yale University Press, 1956.

Spence, K. W., & Spence, J. T. (Eds.) *The Psychology of Learning and Motivation*, 1967, **1**.

Sperling, S. E. Reversal learning and resistance to extinction. A review of the rat literature. *Psychological Bulletin*, 1965, **63**, 281-297. (a)

Sperling, S. E. Reversal learning and resistance to extinction: A supplementary report. *Psychological Bulletin*, 1965, **64**, 310-312. (b)

Spitz, H. H. Cortical satiation as a common factor in perception and abstraction: some postulated relationships based on the performance of atypical groups. *American Journal of Mental Deficiency*, 1959, **63**, 633-638.

Spitz, H. H. Field theory in mental deficiency. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 11-40.

Spitz, H. H. The role of input organization in the learning and memory of mental retardates. In N. R. Ellis (Ed.), *International Review of Research on Mental Retardation*, 1966, **2**, 29-56.

Spitz, H. H. Effects of stimulus information reduction on the search time of retarded adolescents and normal children. *Journal of Experimental Psychology*, 1969, **82**, 482-487.

Spitz, H. H., & Blackman, L. S. A comparison of mental retardates and normals on visual figure aftereffects and reversible figures. *Journal of Abnormal and Social Psychology*, 1959, **58**, 105-110.

Spradlin, J. E., & Girardeau, F. L. The behavior of moderately and severely retarded persons. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*, 1966, **1**, 257-298.

Sprague, R. L., & Binder, A. Verbal and motor perceptual responses of the retarded to ambiguous stimuli. *American Journal of Mental Deficiency*, 1966, **71**, 48-54.

Starkman, S. S., & Cromwell, R. L. Self-evaluation of performance and subsequent practice effects in mental defectives. *Psychological Reports*, 1958, **4**, 414.

Stein, L. Reciprocal action of reward and punishment mechanisms. In R. G. Heath (Ed.), *The role of pleasure in behavior*. New York: Hoeber, 1964. Pp. 113-139.

Stephens, W. E. Category usage of normal and subnormal children on three types of categories. *American Journal of Mental Deficiency*, 1966, **71**, 266-273.

Stevens, H., & Heber, R. F. (Eds.), *Mental retardation*. Chicago, Illinois: University of Chicago Press, 1964.

Stevens, S. S. (Ed.) *Handbook of experimental psychology*. New York: Wiley, 1951.

Stevenson, H. W. Discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 424-438.

Stevenson, H. W., & Iscoe, I. Transposition in the feebleminded. *Journal of Experimental Psychology*, 1955, **49**, 11-15.

Stevenson, H. W., & Odom, R. D. Children's behavior in a probabilistic situation. *Journal of Experimental Psychology*, 1964, **68**, 260-268.

Strauss, A. A., & Lehtinen, L. E. *Psychopathology and education of the brain-injured child*. New York: Greene & Stratton, 1947.

Suboski, M. D. The analysis of classical discrimination conditioning experiments. *Psychological Bulletin*, 1967, **68**, 235-242.

Suppes, P., & Ginsberg, R. A fundamental property of all-or-none models, binomial distribution of responses prior to conditioning, with application to concept formation in children. *Psychological Review*, 1963, **70**, 139-161.

Suppes, P., Jerman, M., & Brian, D. *Computer-assisted instructions: Stanford's 1965-66 arithmetic program.* New York: Academic Press, 1968.

Sutherland, N. S. Stimulus analyzing mechanisms. In *proceedings of the symposium of the mechanization of thought processes.* Vol. II. London: Her Majesty's Stationary Office, 1959. Pp. 575-609.

Taffel, C. Anxiety and the conditioning of verbal behavior. *Journal of Abnormal and Social Psychology*, 1955, **51**, 496-501.

Terrace, H. S. Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 1-27. (a)

Terrace, H. S. Errorless transfer of discrimination across two continua. *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 223-232. (b)

Theios, J. P., & Blosser, D. The overlearning reversal effect and magnitude of reward. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 252-257.

Thorndike, E. L. *Human learning.* New York: Century, 1931.

Thurstone, E. L. Primary Mental Abilities. *Psychometric Monographs*, 1938, No. 1.

Tolman, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.

Touchette, P. E. The effects of graduated stimulus change on the acquisition of a simple discrimination in severely retarded boys. *Journal of the Experimental Analysis of Behavior*, 1968, **11**, 39-48.

Trabasso, T. Stimulus emphasis and all-or-none learning of concept identification. *Journal of Experimental Psychology*, 1963, **65**, 395-406.

Trabasso, T., & Bower, G. H. *Attention in learning: Theory and research.* New York: Wiley, 1968.

Tryon, R. C. Genetic differences in maze-learning ability in rats. *Yearbook of the National Society for the Study of Education*, 1940, **1**, 111-119.

Tulving, E. Subjective organization in free recall of "unrelated" words. *Psychological Review*, 1962, **69**, 344-354.

Turnure, J., & Zigler, E. Outer-directedness in the problem solving of normal and retarded children. *Journal of Abnormal and Social Psychology*, 1964, **69**, 427-436.

Turrisi, F. D., Eimas, P. D., & Shepp, B. E. Extradimensional and interdimensional shifts under constant irrelevant training conditions in the rat. Unpublished manuscript on file at Biological Science Library, Brown University, 1965.

Underwood, B. J., & Schultz, R. W. *Meaningfulness and verbal learning.* Philadelphia, Pennsylvania: Lippincott, 1960.

Vergason, G. A. Retention in retarded and normal subjects as a function of original training. *American Journal of Mental Deficiency*, 1964, **68**, 623-629.

Voeks, V. W. Acquisition of S-R connections: a test of Hull's and Guthrie's theories. *Journal of Experimental Psychology*, 1954, **47**, 137-147.

Wallace, W. P., & Underwood, B. J. Implicit responses and the role of intralist similarity in verbal learning by normal and retarded subjects. *Journal of Educational Psychology*, 1964, **55**, 362-370.

Warren, J. M., & Warren, H. B. Spatial reversal learning by kittens. *Journal of Comparative and Physiological Psychology*, 1967, **63**, 523-525.

Waugh, N. C., & Norman, D. A. Primary Memory. *Psychological Review*, 1965, **72**, 89-104.

Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *Journal of Experimental Psychology*, 1958, **56**, 151-158.

Weir, M. W. Developmental changes in problem-solving strategy. *Psychological Review*, 1964, **71**, 473-490.

Weir, M. W., & Stevenson, H. W. The effect of verbalization in children's learning as a function of a chronological age. *Child Development*, 1959, **30**, 143-149.

Weisberg, P. Social and nonsocial conditioning of infant vocalization. *Child Development*, 1963, **34**, 377-388.

White, S. H. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in Child Development and Behavior*. 1965, **2**, 187-220. (a)

White, S. H. Tranining and timing in the generalization of a voluntary response. *Journal of Experimental Psychology*, 1965, **69**, 269-275. (b)

White, S. H. Age differences in reaction to stimulus variation. In O. J. Harvey (Ed.), *Experience, structure, and adaptability*. New York: Springer, 1966. Pp. 95-122.

Whitty, C. W., & Zangwell, O. L. (Eds.), *Amnesia*. London and Washington, D. C.: Butterworth, 1966.

Wickelgren, W. A. Sparing of short-term memory in an amnesic potent: Implications for strength theory of memory. *Neuropsychologia*, 1968, **6**, 235-244.

Witryol, S. L., Lowden, L. M., & Fagan, J. F. Incentive effects upon attention in children's discrimination learning. *Journal of Experimental Child Psychology*, 1967, **5**, 94-108.

Woodward, W. M. The application of Piaget's concepts to mental deficiency. In *Proceedings of Conference on the Scientific Study of Mental Deficiency, London, 1960*, 1962, **2**, 437-443.

Wyckoff, L. B., Jr. The role of observing responses in discrimination learning. Part I. *Psychological Review*, 1952, **59**, 431-442.

Zeaman, D., & House, B. J. The growth and decay of reactive inhibition as measured by alternation behavior. *Journal of Experimental Psychology*, 1951, **41**, 177-186.

Zeaman, D., & House, B. J. Approach and avoidance in the discrimination learning of retardates. *Child Development*, 1962, **33**, 355-372.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 159-223.

Zeaman, D., & House, B. J. The relation of IQ and learning. In R. M. Gangné (Ed.), *Learning and individual differences*. Columbia, Ohio: Merrill, Books 1967. Pp. 192-212.

Zeaman, D., House, B. J., & Orlando, R. Use of special training conditions in visual discrimination learning with imbeciles. *American Journal of Mental Deficiency*, 1958, **63**, 453-459.

Zeaman, D., & Kaufman, H. Individual differences and theory in a motor learning task. *Psychological Monographs*, 1955, **69**, No. 391.

Zigler, E. An overview of research in learning, motivation, and perception. *Exceptional Children*, 1962, **28**, 9, 455-458.

Zigler, E. Familial mental retardation: A continuing dilemma. *Science*, 1967, **155**, 292-298.

Zigler, E. Developmental versus difference theories of mental retardation and the problem of motivation. *American Journal of Mental Deficiency*, 1969, **73**, 536-556.

Author Index

Numbers in italics refer to the pages on which the complete references are listed.

A

Achenbach, T., 69, *189*
Agranoff, B. W., 33, *192*
Allen, G. A., 92, *189*
Altman, L., 81, *190*
Asratyan, E., 61, *189*
Astrup, C., 58, *189*
Atkinson, J. W., 40, *199*
Atkinson, R. C., 45, 115, 116, 163, 184, *189, 196, 201*

B

Baer, D. M., 139, *190, 200*
Bailer, I., *190*
Bandura, A., 38, *189*
Barber, L., 91, *193*
Barnett, C. D., 62, 68, 80, 104, 106, 107, 111, *189, 193*
Baron, M. R., 106, *191*
Barrett, B. H., 78, *189*
Bartlett, F., 38, *197*
Bauer, F. J., 83, *189*
Beach, D. R., 50, *194*
Beck, C. H., 178, *189*
Belmont, J. M., 68, *189*
Bensberg, G. J., 139, *189*
Berkowitz, L., *189*
Berkson, G., 112, *190*
Bevan, W., *195*, 196
Bialer, I., 128, 129, *190*
Bijou, S. W., 62, 91, 139, 140, 141, *190, 200*
Bilodeau, E. A., 106, *191*
Binder, A., 78, 85, *190, 203*
Birnbrauer, J. S., 62, 91, *190*
Blackman, L. S., 34, 91, 133, *190, 203*
Blake, K. A., 63, *192*
Blank, J. P., 108, *190*
Blank, M., 81, *190*
Blehert, S. L., 163, *190*
Bliss, E. L., *190, 196*
Blosser, D., 178, *204*
Blue, C. M., 114, *190*
Bousfield, W. A., 48, 134, *190*
Bovet, D., 46, *191*
Bovet-Nitti, F., 46, *191*
Bower, G. H., 1, 4, 10, 33, 37, 76, 81, 146, *189, 191, 196, 202, 204*
Brackbill, Y., 46, *191*
Brian, D., 93, 145, *204*
Bright, P. I., 33, *192*
Broadbent, D. E., 115, *191*
Brown, J. S., 106, *191*
Brown, Z., 175, *191*
Bryant, P. E., 62, 86, *191, 200*
Bush, R. R., 154, *191*

C

Caldwell, E. C., 88, *191*
Calfee, R. C., 184, *191*
Campbell, B. A., *191, 193*
Campione, J., 170, 178, *191*
Cantor, G. N., 80, 104, 106, 107, 108, 109, *189, 191*
Capaldi, E. J., 15, *191*
Caron, A. J., 81, *191*
Chesler, P., 38, *197*
Chinsky, J. M., 50, *194*
Church, R. M., *191, 193*
Clark, R. A., 40, *199*
Cofer, C. N., *191, 201*
Cohen, L. B., 83, *196*
Cole, M., 48, 49, *191, 202*
Cook, S. W., 77, *191*
Craig, G. J., 48, *192*
Cromwell, R. L., 58, 126, 127, 128, 129, 130, 131, *190, 192, 203*
Cruickshank, W. M., 63, *192*
Cunningham, M. A., 82, *192*

D

D'Amato, M. R., 178, *192*
Dameron, L. E., 58, *195*
Davis, R. E., 33, *192*
Denny, M. R., 2, 52, 56, 57, 58, 59, 61, 63, 65, 66, 67, 68, 71, 121, 122, 123, 125, 167, *192*
Dickerson, D. J., 82, 170, *192*
Distefano, M., 64, *193*
Doubros, S. G., 58, *192*

E

Egger, M. D., 10, *192*
Eimas, P. D., 170, 176, 178, *192*, *202*, *204*
Eisman, B. S., 67, *192*
Ellis, N. R., 2, 33, 58, 62, 63, 64, 68, 69, 104, 107, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 180, *189*, *192*, *193*, *194*, *195*, *197*, *198*, *201*, *202*, *203*
Ellson, D. G., 91, *193*
Erlenmeyer-Kimling, L., 45, *196*
Ervin, F. R., 75, *194*
Esterson, J., 48, *190*
Estes, W. K., 4, 8, 9, 11, 24, 26, 40, 46, 61, 74, 75, 92, 123, 125, 146, 154, 157, 175, *189*, *193*, *199*, *200*

F

Fagan J. F., 81, *205*
Fernald, G. M., 168, *193*
Ferster, C. B., 138, *193*
Fields, P. E., 82, *194*
Fishkin, S., 85, *202*
Flavell, J. H., 49, 50, *194*, *200*
Fletcher, F. M., 80, *200*
Foshee, J. G., 58, 107, *192*, *202*
Frankel, F., 48, 49, *191*
Franks, C. M., 58, *194*
Franks, V., 58, *194*
Freitas, L., 78, *199*
Fuchs, J. E., 111, *200*
Fuller, J. L., 45, 46, *194*
Fuller, P. R., 78, *194*
Furth, H. G., 48, 178, *194*, *199*

G

Gagnê, R. M., 31, 53, *194*, *199*
Garcia, J., 75, *194*
Gardner, W. I., 128, *194*
Gewirtz, J. L., 131, *194*
Gibson, E. J., 35, *194*
Gilmore, A. S., 91, *199*
Ginsberg, R., 149, *203*
Girardeau, F. L., 61, 62, 82, 111, *192*, *194*, *203*
Glaser, R., 3, 89, *194*, *199*
Gormezano, I., 72, *194*
Goslin, D. A., *194*
Goulet, L. R., 48, *194*
Grant, D. A., 72, *194*
Gray, S. W., 141, *195*
Greenberg, R., 26, *195*
Greeno, J. G., 10, *195*
Griffith, B. C., 68, *195*
Grings, W. W., 58, *195*
Grossman, S. P., 19, *195*
Grusec, J. E., 38, *189*
Grusec, T., 81, *191*
Guilani, B., 78, *199*
Guilford, J. P., 1, 8, *195*
Guthrie, E. R., 7, 8, 72, 121, 123, *195*

H

Hagen, J. W., 49, *197*
Hall, J. F., 71, 74, *195*
Hall, V. C., 88, *191*
Hamilton, J., 68, *195*
Hardy, M. W., 91, *199*
Harlow, H. F., 26, 47, 66, 147, *195*
Harris, J. D., 73, *195*
Harris, P., 91, *193*
Harris, R. E., 77, *191*
Harvey, O. J., *195*, *200*, *205*
Hawles, T. G., 49, *200*
Hayes, K. J., 149, *195*
Haywood, C., *195*
Haywood, H. C., 36, 141, *195*, *201*
Heal, L. W., 65, 179, *201*
Heath, R. G., *195*
Hebb, D. O., 32, 35, 36, 110, *195*
Heber, R. F., 2, 128, *195*, *203*

Helson, H., *195*, *196*
Hermelin, B., 112, *190*
Heron, W. T., 46, *195*
Hilgard, E. R., 1, 45, 77, *195*, *196*
Hill, S. D., 47, *196*
Hirsch, J., 3, 45, *196*
Honig, W. K., *196*
Hope, R., 117, *193*
House, B. J., 24, 26, 27, 47, 64, 65, 66, 79, 80, 82, 106, 115, 125, 143, 146, 148, 149, 151, 154, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 171, 174, 175, 176, 178, 180, 181, *196*, *205*
Hull, C. L., 7, 8, 26, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 108, 109, 110, 111, 113, 115, 121, 122, 126, 137, 146, 148, 149, 150, 173, *196*
Hunter, W. S., 111, *199*
Hyman, L., 170, 178, *191*

I

Irion, A. L., 71, *199*
Irwin, O. C., 46, *196*
Isaacson, R. L., 76, *200*
Iscoe, I., 66, *203*

J

Jacobs, A., 108, *196*
Jagoda, H., 178, *192*
Jarvik, M. E., 83, *196*
Jeffrey, W. E., 83, *196*
Jenkins, H. M., 83, *196*
Jensen, A. R., 68, 86, 135, *196*, *197*
Jerman, M., 93, 145, *204*
John, E. R., 38, *197*
Johnson, B. M., 62, *197*
Jones, M. R., *197*, *200*
Jones, R. W., 63, 107, *197*

K

Kagan, J., *197*
Kamin, L. J., 33, 72, 73, 74, *197*
Kaufman, H., 4, *205*
Kaufman, M. E., 66, *197*
Kendler, H. H., 50, 154, 171, 173, 179, 181, 185, *197*
Kendler, T. S., 50, 154, 171, 173, 179, 181, 185, *197*
Keppel, G., 48, *197*
Kidder, J. D., 62, 91, *190*
Kimble, G. A., 72, 77, *197*
Kingsley, P. R., 49, *197*
Klaus, R. A., 141, *195*
Koch, S., *193*, *195*, *197*, *198*, *199*, *200*
Koelling, R. A., 75, *194*
Kohler, W., 110, 133, *197*
Koski, C. H., 59, 60, *201*
Kounin, J., 132, *197*
Krasner, L., 22, 61, *190*, *197*, *198*
Kuenne, M. R., 50, *198*

L

Lashley, K. S., 82, *198*
Lauer, D. W., 24, 26, *193*
Laurence, M. W., 49, *198*
Lawler, J., 62, *190*
Lawrence, D. H., 10, 80, 81, 83, 189, *198*
Lawson, R., 91, *198*
Lazarovitz, L., 76, *191*
Lehtinen, L. E., 168, *203*
Lenneberg, E. H., 2, *198*
Leontiev, A. N., 88, *198*
Levine, M., 81, *198*
Lewin, K., 126, 132, *198*
Lindsley, O. R., 78, *189*, *198*
Lipman, R. S., 57, 63, 67, 68, *195*, *198*
Lipsitt, L. P., 2, 46, 47, *198*, *205*
Locke, B. J., 77, *198*
Lockhart, R. S., 58, *195*
Lovaas, O. J., 40, 78, *199*
Lovejoy, E., 146, 163, 171, *199*
Lowden, L. M., 81, *205*
Lowell, E. L., 40, *199*
Lumsdaine, A. A., 89, *199*
Luria, A. R., 59, *199*

M

McClearn, G. E., 45, *199*
McClelland, D. C., 40, *199*
McConnell, D., 78, *190*
MacCorquodale, K., 9, *193*, *199*
McCrary, J. W., 111, *199*

McGeoch, J. A., 71, *199*
McGowan, B. K., 75, *194*
Mackintosh, N. J., 178, *196*, *199*
Mahler, W. A., 92, *189*
Malpass, L. F., 91, *199*
Maltzman, I., *191*, *202*
Marken, R. S., 185, *197*
Marquis, D. P., 46, *199*
Means, J. R., 85, *202*
Meehl, P., 9, *193*, *199*
Melton, A. W., 4, *193*, *199*, *200*
Menlove, F. L., 38, *189*
Meredith, W., 45, *199*
Metz, J. R., 78, *199*
Meyer, D. R., 26, *199*
Miles, R. C., 26, *199*
Milgram, N. A., 48, 68, 82, 86, *199*
Miller, G. A., 135, *199*
Miller, J., 77, *199*
Miller, M. B., 129, *200*
Miller, N. E., 7, 10, 40, 99, 126, *192*, *200*
Miller, R. E., 83, *200*
Milner, B., 31, *200*
Moely, B. E., 49, *200*
Montessorr, M., 168, *200*
Moore, J. W., 72, *194*
Moore, O. K., 92, *200*
Mosteller, F., 154, *191*
Mueller, C. G., Jr., *193*, *199*
Muenzinger, K. F., 80, *200*
Murphy, J. V., 83, *200*
Musgrave, B., *191*, *201*
Myers, J. L., 48, *192*

N

Neimark, E. D., 9, 74, 146, 157, *200*
Nelson, K., 78, *199*
Noble, C. E., 111, *200*
Noce, J. S., 82, *199*
Norman, D. A., 115, *204*

O

Oatley, K. G., 62, *200*
O'Connor, N., 112, *190*
Odom, R. D., 48, *203*
Ohlick, E. S., 66
Olds, J., 11, *200*
Oliverio, A., 46, *191*
Olson, F. A., 49, *200*
Olton, D. S., 76, *200*
Orlando, R., 61, 62, 65, 79, 82, *190*, *200*, *205*

P

Palk, B. E., 58, *192*
Paraskevopoulos, I., 49, *200*
Passey, G. E., 72, *201*
Pavlov, I. P., 72, 75, 110, 121, *201*
Pennington, D. F., 25, *201*
Peterson, L. R., 115, *201*
Peterson, W. M., 66, *197*
Phillips, J. L., 116, *201*
Piaget, J., 2, 32, *201*
Postman, L., 9, *201*
Pressey, S. L., 89, *201*
Price, J. E., 91, *201*
Prokasy, W. F., *197*, *201*
Pryer, M. W., 63, 64, 68, 104, 107, 111, *189*, *193*
Pryer, R. S., 64, *193*

R

Ramsey, O. E., 78, *201*
Rescorla, R. A., 73, *201*
Restle, F., 170, *201*
Ringelheim, D., 127, *201*
Rohwer, W. D., Jr., 86, 135, *197*
Ross, L. E., 20, 59, 60, 65, 66, 179, *201*
Rotter, J. B., 126, 130, *201*
Rudel, R. G., 66, *201*

S

Sainsbury, R. S., 83, *196*
Sanders, B., 65, 179, *201*
Schaeffer, B., 40, *199*
Schneider, D. E., 72, *194*
Schoeffler, M. S., 125, *193*
Schoenfeld, W. N., Jr., *193*, *199*
Schultz, R. W., 85, *204*
Scott, K. G., 4, *201*
Sersen, E. A., 58, *189*
Sharp, D., 48, 49, *191*

Sheer, D. E., *200, 202*
Shepp, B. E., 83, 167, 169, 170, 176, 180, *202, 204*
Shif, Z. I., 92, *202*
Shiffrin, R. M., 115, 116, 184, 189, *201*
Sidman, M., 139, 141, 142, 143, *202*
Siegel, P. S., 107, *202*
Simmons, J. Q., 40, *199*
Simmons, M. W., 47, *202*
Simon, B., *198, 202*
Sjoholm, N. A., 78, *190*
Skinner, B. F., 7, 9, 46, 79, 89, 90, 137, 138, 146, *193, 195, 202*
Slavina, L. S., 89, *202*
Slerner, 178, *189*
Sloan, W., 63, 69, *193*
Smith, J., 175, *191*
Smith, M., 27, *196*
Smith, M. P., 85, 91, *190, 202*
Solomon, R. L., 9, *202*
Spearman, C., 53, *202*
Spence, J. T., *189, 191, 202, 203*
Spence, K. W., 8, 26, 97, 100, 104, 105, 106, 108, 126, 146, 148, 149, 150, 170, 173, *191, 202, 203*
Sperling, S. E., 178, *203*
Spiker, C. C., 2, *198, 205*
Spitz, H. H., 34, 68, 133, 134, 135, *195, 203*
Spradlin, J. E., 61, 62, 82, *192, 194, 203*
Sprague, R. L., 85, *203*
Starkman, S. S., 127, *203*
Starr, R., 76, *191*
Stein, L., 40, *203*
Stephens, W. E., 68, *203*
Stevens, H., *203*
Stevens, S. S., *202, 203*
Stevenson, H. W., 48, 57, 64, 66, 168, *203, 205*
Stoddard, L. T., 139, 141, 142, 143, *202*
Strauss, A. A., 168, *203*
Suboski, M. D., 174, *203*
Suppes, P., 93, 145, 149, *203, 204*
Sutherland, N. S., 154, 171, 175, *204*

T

Taffel, C., 22, *204*
Tapp, J. T., 36, 141, *193, 195*
Taylor, J., 109
Terrace, H. C., 79, 80, 90, 122, *204*
Terrell, C. G., 113
Theios, J. P., 178, *204*
Thompson, S. R., 25, *201*
Thompson, W. R., 45, *194*
Thorndike, E. L., 4, *204*
Thurstone, E. L., 53, *204*
Tinson, C., 62, *200*
Tolman, E. C., 9, 126, *204*
Touchette, R. E., 180, *204*
Trabasso, T., 4, 10, 33, 146, 150, *204*
Tryon, R. C., 46, *204*
Tulving, E., 134, *204*
Turner, L. H., 9, *202*
Turnure, J., 69, *204*
Turrisi, F. D., 169, 170, 176, *202, 204*

U

Underwood, B. J., 26, 69, 85, *195, 204*
Ullmann, L., *190, 198*

V

Vergason, G. A., 114, *204*
Verplanck, W. S., *193, 199*
Victor, I., 38, *197*
Vocks, V. W., 18, 74, *204*

W

Wallace, W. P., 69, *204*
Wallach, H., 133, *197*
Warren, H. B., 178, *204*
Warren, J. M., 178, *189, 204*
Watson, L. J., 91, *198*
Waugh, N. C., 115, *204*
Weinstock, S., 25, *204*
Weir, M. W., 48, 168, *205*
Weisberg, P., 46, *205*
Whalen, C., 78, *199*
White, S. H., 50, 51, *205*
Whitmarsh, G. A., 48, *190*
Whitty, C. W., *200, 205*
Wickelgren, W. A., 31, *205*
Williams, C. F., 91, *199*

Witryol, S. L., 81, *205*
Wolf, M. M., 62, 91, *190*
Woodward, W. M., 32, *205*
Wortis, J., 58, *189*
Wright, J. C., *197*
Wyckoft, L. B., Jr., 170, *205*

Y

Yager, J., 59, 60, *201*
Youniss, J., 178, *194*

Z

Zangwell, O. L., *200*, *205*
Zeaman, D., 4, 24, 26, 27, 47, 64, 65, 66, 79, 80, 82, 106, 115, 125, 143, 146, 148, 149, 150, 151, 154, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 174, 175, 176, 178, 180, 181, *191*, *196*, *202*, *205*
Zigler, E., 69, 185, *189*, *204*, *205*
Zimmerman, D. W., 113

Subject Index

A

Achievement motives, 40
Adaptation-to-cues theory, 170
All-or-none learning, 35
  backward learning curves for, 150
Alpha rhythm blocking, 112-113
Anticipation of reward and punishment, 7, 11, 14, 20-21, 23-24, 38, 40, 72-74
  conditioned, 9
  level of, 24
  variations in, 14
Anticipations of UR, 17
Anxiety, 109
Appetitive conditioning, 46, 58-59
  difference in rate of acquisition of, 58-59
  in infancy, 46
Arithmetic learning, 89, 91, 93
Association theory, 7, 51, 78, 121, 137, 146, 154
Associative connections
  random fluctuations of, 12, 17, 35
  in relation to intelligence, 4, 166
Associative elements, overlapping, 9
Associative processes, 33-34
Associative relations, 8, 16
  hierarchical organization of, 8
  transfer of, 16
Attending response, 16, 18, 21, 29, 60, 81-82, 84, 115, 152-169, 175-176
  in classical conditioning, 18
  in discrimination learning, 23, 25, 164, 175-176
  effect of environment on, 168
    of instruction on, 81-82, 168
  vs. instrumental response, rate of learning, 175-176
  learning of, 26, 166
  in operant conditioning, 21-22
  in relation to intelligence, 48, 106, 122, 166
  transfer to relevant dimension of, 23-24, 26-27
Attention, deficit in, 122
Attentional theory, 106, 146-166, 170-171, 181
  assumptions of, 151-156
  evaluation of, 168-169, 173-174, 176, 180-182
  simplifying assumptions for, 154
"Autotelic environment," 92
Aversive conditioning, 9, 46, 58-59
Avoidance conditioning, 46, 76, 91
  "safety signal" in, 76
  selective breeding and rate of, 46

B

Backward learning curves, 47, 65, 149-151, 153, 159, 161-162, 164, 169
  for all-or-none learning, 150
  for computer simulation subjects, 150, 160
  families of, 160, 169
  form of, 149-150, 159
  initial plateau of, 164-165, 166-167
  for problems of different difficulty, 151
  for shift behavior, 164-165
"Barrier" to prevent incorrect responses, 122
  effect on learning of, 125
  fading of, 125
Behavior genetics, 45
Behavioral event, 8, 9, 12, 14, 23
"Blocking" of conditioning, 33, 73-74

C

Capacity, 30-31
  defined, 30
  development of, 31
  genetic determiners of, 31
  limits of, 31
Categorizing responses, 15
Categorizing strategies, 50
"Central nervous system integrity," 111
Choice point, exposure to stimuli at, 16, 24, 26, 41, 80-81
Classical aversive conditioning, 46, 58-59

Classical conditioning, 17-20, 58-61, 71-77
  advantages in studies of learning and mental development, 57, 60
  attending responses in, 18, 72
  comparative studies of, 58-61
  contextual stimuli in, 59-60, 74
  correlation of intelligence and acquisition rate of, 58
  cortical function in, 59
  differences
    in acquisition rate in, 58-59
    of operant conditioning from, 61
  effect of previous learning on, 56-57, 60
  of eyelid response, 17, 58-59
  individual differences in rate of, 18, 19
  magnitude of UR in, 18, 72
  negative drive amplifier in, 17
  partial reinforcement effect in, 59-60
  problems of comparative study of, 60-61
  rate of acquisition of, 58-59, 71-77
    of extinction of, 59-60
  repeated acquisition and extinction in, 59-60
  representation of stimuli in memory in, 17-18
  role of selective attention in, 174-175
  verbal rehearsal in, 20
Classical conditioning experiment
  comparability of procedures used with various organisms in, 60
  controls in, 18, 60-61
Clustering, 15, 48-49, 134-136
  effect of instructions on, 49, 135-136
    of repetition on, 134-135
Clusters of behavior, reinforcement of, 127
Coding, cross modal, 10, 27
Coding strategy, 27, 31
  long-term recall and, 27
Comparative studies of retardates and normals, 52-70
  aim of, 54
  in classical conditioning, 58-61
  in complex verbal processes, 68-69
  computers in, 54
  control groups in, 53, 55-57
  in discrimination learning, 64-66
  in learning set, 66-67
  methodology in, 52-57
  multiple control groups in, 55, 57
  in operant conditioning, 61-62
  in perceptual-motor learning, 62-64
  in problem solving, 69
  in reversal shift behavior, 65-66
  in transposition, 66
  in use of verbal mediators, 68-69
  in verbal learning, 67-68
Computer analogy, 7
Computer-assisted instruction, 54, 93, 145
Computer simulation
  backward learning curves from, 150-159, 160, 165
  to test attention model, 159-160
Concept learning, 14, 16, 48, 68
Conditioned anticipation, 9, 75
  of aversive events, 9
  of food, 75
  of reward, 9
Conditioned discrimination, 17-18, 20, 23, 174
Conditioned emotional responses (CER), 11, 73, 75, 140
Conditioned responses, 19, 75-76
  blocking of, 33, 73-74
  consolidation of, 33
  rate of acquisition of, 18, 19, 32, 33, 75
Conditioning, *see also* Classical conditioning, Operant conditioning
  under curare, 9, 72
  eyelid, 17-20, 174-175
  in infancy, 46-47
  invariance of learning rate parameters in, 167
  in monkeys, 47
  one-trial, 35, 36, 74
  salivary, 17
  selective breeding and rate of, 46
  temporal, 110
  trace, 110
Consolidation of long term storage, *see* Transfer to long term storage
Contextual cues
  in classical conditioning, 18-19, 59-60, 74

  relation to mental ability to, 19
Contiguity, association by, 8, 9, 12, 23, 121, 123
Contiguity theory, 121
Cortical satiation, 133
Counter conditioning, 18-19, 121
Covert scanning of alternatives, 41
"Crutch" cues, 122, 124-125
CS-US interval
  in avoidance conditioning, 76
  effect of nature of CS on, 75
  length of, 74
  probability of conditioning and, 74
  in relation to CR, 75
Cue-learning strategy, 69
"Cumulative-learning" hypothesis, 31-32
Custodial institutional residence, effect on development of, 141, 187

D

Deficit in attention, 122
Delay of reward, 25, 108, 129
Delayed reaction, 59, 110, 111, 122
  of brain-damaged subjects, 112
  intelligence and, 113
  phylogenetic level and, 112
  rehearsal strategies and, 112, 113
  stimulus trace theory of, 112-113
Denny's elicitation theory, 121-124
  contiguity in, 121
  reinforcement in, 121
  role of context in, 123
  stimulus sampling in, 122, 123
  transfer in, 123-124
Development, deviation from normal, 2
Developmental theories, 2
Developmental trends in learning, 47, 50-51
Differentiation, 132, 138
Dimensions, transfer along, 16, 171
Discrimination, 47-48, 138, 141-143
  in classical conditioning, 17-18, 20, 23, 174
  of complex patterns, 82-83
  concurrent, 82-83
  effect of qualitatively different rewards on, 180
  of geometric forms, 142-143
  see also Reversal shift
Discrimination behavior, restricted repertoire of, 139-140
Discrimination learning, 14, 23-26, 50-51, 64-66, 79-84, 146-182, 188
  asymptotes in, 166-167
  attending responses in, 23-24, 64, 65, 80, 81-83, 146, 151, 159-160, 174
  effect of distinctive (novel) cues on, 83, 165-166
    of instruction on, 82
    of intertrial interval on, 25
    of motivation on, 65, 78-79
    of naming stimuli on, 82
    of previous experience on, 24, 64-65, 80-81
    of rehearsal on, 26
    of spatial contiguity on, 82
    of temporal contiguity on, 82-83
  errorless, 79-80
  Hull-Spence explanation of, 148
  individual differences in rate of, 24
  inhibition deficit in, 65
  intelligence and, 64-65
  invariance of learning rate parameters in, 166-167
  latency of response in, 51
  optional shift studies of, 50
  pretraining for, 79-80, 81, 88
  rate of, 24, 47, 78-81, 83, 102
  relationship to conditioning of, 23
  relative difficulty of object and pattern problems, 161-162, 168, 180
  role of short-term memory in, 25
  task difficulty in, 64-65, 80, 159, 161, 167
  transfer in, 50, 80, 81
  two-stage process of, 150, 151-160, 170
Discrimination learning set, *see* Learning set
Discrimination shift, 169-180
  analysis of transfer in, 177
  attention theory and, 169-171
  cross-modal, 81, 175
  mediation theory and, 171-173
  nonreversal, 50
  optional, 50

  overtraining and, 170, 175-176, 178-179
  reversal, 25, 48, 50, 65, 66, 177
  speed of intra- and extradimensional, 169-170, 172, 175
  variation of irrelevant cue in, 176
  see also Reversal shift
Discrimination theory, 26-27, 143-144, 146-166, 170, 180-182
Distinctive features of stimuli, 85-86
Distributed practice, 25, 107, 123
Drive, 10-11, 98, 102-103, 105, 108
  effect on stages of learning on, 103
Drive amplifier, 17
Drive amplifier sources, 72-73

E

Easy-to-hard transfer, 64-68, 80, 124-125, 159, 161, 168-169
Educational situation
  instructional methods in, 80-87, 122-125, 168-180, 186-187
  technologies in, 87, 89-93, 125
Electrical stimulation of midbrain centers, 10
Electroconvulsive shock, 9
Elicitation hierarchy, 121-124
  motivation in, 124
  satiation in, 121
Ellis's stimulus trace hypothesis, 110-114
Emotional disposition, 39-41, 54-55
Environment, 168
  early, 39-40, 141
  impoverished, 35-36, 132-133, 186-187
Errorless discrimination learning, 79-80, 90, 142-143
Errors, diagnostic use of, 91
Excitatory potential, 97-100, 103
Excitatory tendency, 98-100
Expectation
  of reinforcement, 126-127
    generalization of, 127
  of success, 24, 127-128
    effect of reward on, 127-128
    rate of development of, 128
Extinctions of CR
  by counterconditioning, 18-19, 121
  prior to first unreinforced trial, 77
  rate of, 18, 58-61, 61-62
  by retardates, 59-60
Extradimensional shift, 163-165, 169-170, 172, 175-178
  attention theory and, 163-165
  mediation theory and, 172

F

Fading, 143
  of barriers preventing errors, 125
Fading techniques, 80, 90, 122, 124-125
  in programmed instruction, 90
Feedback, 14, 71
  differential, 123
  in discrimination learning, 11-12, 23
  extension to other stimuli of, 11
  facilitatory, 11, 21, 24, 39
  inhibitory, 11, 14, 39
  positive, 14
  preferential selection of positive, 11
  suppression of outgoing behaviors by, 11
Field theory, 132-136
  learning of retardates in, 134
  postulates of, 133-134
Fluctuation of stimulus context, 12
Forced pause at choice point, 16, 24, 26, 41, 80-81
  in children's discrimination learning, 81
Free recall, 14, 48-49, 69, 134-136, 188
  clustering in, 48-49, 134-135
  developmental studies of, 48-49
  implicit verbal associations and, 69
  imposed clustering in, 135-136
  relation of intelligence to, 135
  spontaneous clustering in, 49

G

Genetic factors in intelligence, 30, 45-46, 140, 183
Gestalt psychology, 7, 110, 132-134

H

Habit strength, 98
  growth of, 98-99, 101-102, 104-105

Habituation, 179-180
Half-reversal shift, 185
Hebb's neurophysiological theory, 32, 110
Hierarchical organization of associations, 8, 51
Hierarchies of S & R units, 14-15
Higher-order behavior units, formation of, 8, 34-36, 38, 51
Hull-Spence theory, 104-109, 148
Hull's theory, 97-104
  discrimination training in, 99-100
  dominance in American learning theory of, 97
  individual differences in, 100-104, 105
  postulate of, 97-100

I

Image, 12
Immediate memory, 18, 31, 115
  see also Short term memory
Implicit associations to verbal stimuli, 69
Inadequate reinforcement history, 139-141
  as cause of retardation, 139-140
  effect of reinforcement procedure on, 140-141
  genetic factors in, 140
  role of environment in, 139
Incentive motivation, 99-100, 103, 105, 108
  effect at different stages of learning of, 103
  individual differences in, 103-104, 108
Incentives, preferences among, 108
Incidental learning, 9
Individual differences, 3, 4, 23-24
  in learning ability, 30, 183
  in rate of conditioning, 18-19, 21
    of learning, 52-53
  in reaction to reinforcement, 129
Information theory, 7
Inhibition, 93, 98, 99-100, 102, 104, 133, 146
  deficit in, 41, 59-60
  definition of, 106-108
  in massed-trial learning, 107-108
  in relation to learning ability, 107
Inhibition of delay, 108
Inhibitory defect, 55-60, 63, 65, 101, 122
  in classical conditioning, 59-60
  cortical function and, 59, 112
  in relation to extinction, 59
Inhibitory potential, 98-100
Initial plateau in backward learning curve, 166-167
Intelligence
  factors in, 1, 53-54
  genetic determiners of, 45
  learning ability and, 1, 52-53
  as learning performance, 53-54
  relation of previous experience to, 52-54
Intensity of CS and US, 72
Interference, 34
Interfering associations, 10, 12
Internalization of behavior, 41-42
Interneural connections, 4
Intertrial stimulus variability, 74
Intradimensional shift, 20-27, 163-164, 169-170, 172, 175-178
  attention theory and, 26-27, 163-165
  mediation theory and, 172
Invariance of learning rate parameters, 166-167
Irrelevant cues, random variation of, 176
Irrelevant dimension
  effect of reward while attending to, 124-125, 155-156
    of verbalization of, 82
  probability of attending to, 152-153, 157
    of correct choice response while attending to, 153, 157

K

Korsakoff's syndrome, 33

L

Labeling strategies, 50, 85-86, 171, 173
Language, development of, 1, 50
Language learning, 16, 28
Latency, relation to correct responding of, 51
Latent learning, 9
Learning

  control of previous opportunities for, 54-55
  defined, 8
  longitudinal study of, 54
  from models, *see* Observational learning
  parameters of, 52-53, 70, 97-98, 149
  performance and, 97
Learning ability, 30-32, 159
  asymptotes of, 101
  correlation among rate constants of, 101-102, 156-157, 166
  development of, 31-42
  developmental studies of, 46-51, 93
  differences in drive and, 102-103
    in inhibitory tendencies and, 102
    in stimulus generalization and, 102
  genetic factors in, 30, 31
  genetic studies of, 45-56
  individual differences in, 30, 31, 54-55, 100-104, 166
  previous experience and, 29-31, 53-55, 57, 183
Learning capacities, development of, 31, 46-51, 56-57, 184
Learning to learn, 15-16, 26-28, 67
Learning rate parameters, 101-102, 156-157, 166-167, 169
  defined, 156-157
  intelligence and, 53-54, 101-102, 166-167
  invariance of, 166-167
  overtraining and, 166-167, 178
Learning to read, 14, 16
Learning set, 15, 26-27, 66-67, 147
  development of, 15-16, 66, 67
  in discrimination theory, 26
  among mental retardates, 66-67
  perseverative errors in, 67
  in stimulus sampling theory, 26
Learning to speak, 14
Learning theory, 2-3
  applied in classroom, 86-87
  individual differences in relation to, 4, 5
  simplifying assumptions in, 4
Linguistic skills, teaching of, 92
"Locus of control," 129
Long-term memory, 20, 28, 31, 33, 109, 115-116, 117-120
  age and, 33
  all-or-none transfer to, 33
  coding strategies in, 27
  consolidation in, 9, 20, 23, 31-33
  effect of drugs on, 34
  intelligence and, 20, 68, 110
  neurological mechanisms in, 31, 33
  retrievability from, 10
  retrieval strategies from, 49
  transfer to, 9-10, 15, 19, 20, 31-34

M

Massed practice, 63, 85, 107-108, 123
Maturational factors, 1
Maze learning, rate of
  intelligence and, 63-64
  selective breeding and, 45-46
Meaningfulness in paired-associate learning, 85
Mediating response hypothesis, 170-174, 179, 181
  chained responses in, 171
  language development and, 172
  shift studies and, 40, 50, 172
Mediating responses, 15, 48, 50, 170-173
  as attention directors, 173
  developmental stages in, 179
  in facilitating memory storage, 173
  feedback from, 171-173
  function of, 172-173
  labels as, 171-173
  selective association with relevant cues of, 173
  spontaneous, 86
Memory
  disorders of, 31
  immediate, 31
  "temporal stacking" in, 51
Memory trace
  in Gestalt psychology, 110
  two-stage theory of, 46
Midbrain centers, electrical stimulation to, 10-11
Motivation, 10-12, 61, 68-79, 89-90, 97, 133
  response evocation and, 10-12, 99

Motivational systems, development of, 31, 39-41
Motor skills, learning rate parameters in, 167
Music training, 88-89

N

Naming cues, 82
Negative drive amplifier, 17
Neurophysiological mechanisms, development of, 36
Normal control groups, 55-57
Novel stimuli, 83, 165-166, 176
Novelty as stimulus dimension, 166

O

Observational learning, 37-39
Observing response, 16, 152, 164, 170-172, 177, *see also* Attending response
Oddity discrimination problems, 47, 69
  conditional, 47
  developmental study of, 47
  intelligence and, 69
One-trial learning, 35, 36, 65, 74
Opportunities to learn
  confounding of learning set with, 57
  control of, 54-55
  equating of, 56-57
Operant conditioning, 20-23, 61-62
  acquisition procedures for, 61, 77
  attending responses in, 21
  building complex behavior by, 78
  difference from classical conditioning of, 61
  of discrimination, 47, 62, 141-142
  effect of instructions on, 78
    of reward in, 20-21, 61-62, 77-78
  individual differences in rate of, 21
  in infants, 46-47
  partial reinforcement in, 62
  personality factors in, 78
  in practical training, 62, 78, 87
  reinforcement contingencies in, 22, 61
  reinforcement schedules and, 61-62, 77
  role of attention in, 77
Operant conditioning approach, 62, 137-145, 186-187
  applied to individual cases, 87
Operant conditioning experiments, objectives of, 61
Operant discrimination in infancy, 46-47
Organizational strategies, 16, 36, 134-136
Orienting response, *see* Attending response
Overlapping associative elements, 9
Overlearning reversal effect, 66, 178-180
  habituation decrement in, 179
  effect of positional discriminations in, 178-179
  mediation theory and, 180
  reward magnitude and, 178

P

Paired-associate learning, 27, 28-29, 48, 84-86, 135
  acquisition rate in, 32, 67-68, 84-85
  attending responses in, 29, 84
  classical conditioning and, 28
  errors in, 27
  implicit associations in, 69, 84-85
  individual differences in rate of, 28-29
  intelligence and, 29, 67-68
  intertrial interval in, 114
  learning to learn and, 26-28
  meaningfulness and, 85, 135
  mediators in, 28, 85-86
  minimum set of stimulus elements in, 29
  motor responses in, 85
  overlearning in, 114
  rate of, 28-29, 84-86
  rehearsal in, 29, 84
  stimulus selection in, 33
  stimulus trace theory of, 114
  task difficulty in, 68
Pathological cases of retardation, 185-186
Pavlovian conditioning, *see* Classical conditioning
Perception differentiation of, 132-133
Perceptual-motor learning, 22-23, 62-64
Period of "slow learning," 35-36

Perseverative behavior, 132
Perseverative errors, 67
Position habits, 47-48
  reversal of, 48
Positional discrimination, speed of learning, 106, 178-179
Primary memory, *see also* Short-term memory
  in relation to intelligence, 115-119
Proactive inhibition, 27, 119-120
Probability learning, 13-14, 48
  alternation tendencies in, 48
  developmental studies of, 48
  probability matching in, 48
Problem solving
  imitation in, 69
  strategies of retardates in, 69
Programmed instruction, 90, 180-181, *see also* Teaching machines
  fading technique in, 90
Punishment
  in early environment, 39-40
  effect of, 140
    on positive drive system of, 40
  theory of, 39-40
Pursuit-rotor performance, related to intelligence, 63

R

Random fluctuation of associative connections, 12, 17
Rate of classical conditioning, 56-60, 72-77
  during acquisition, 56-57, 58-59
  CA and, 132
  effect of CS-alone trials on, 73
    of CS-US interval on, 74-75
    of experimental controls on, 71
    of frequency of US in absence of CS on, 73
    of instructions on, 77
    of intensity of US on, 72
  during extinction, 59-60
  learning ability and, 32
  redundant stimuli on, 73-75
  stimulus variability on, 74
  modality of CS and US in, 75-76
Rate of discrimination learning, 24, 78-81, 83, 102
Rate of learning
  modification of, 56
  past experience and, 31, 53-56
  qualitative differences in, 69-70
  quantitative differences in, 70
  in relation to selective attention, 70
  spontaneous verbal mediators and, 70
Rate of paired-associate learning
  attending responses and, 84
  CA and, 32
  meaningfulness and, 85
  naming of stimuli and, 85
  rehearsal and, 84
  response interrelationships and, 84
  stimulus patterning and, 84
  stimulus-response associations and, 84-86
  verbal mediators in, 85-86
Rate parameters, *see* Learning rate parameters
Reaction time, 113-114, 128
Reading readiness, 35, 36, 88
  concept of similarity in, 88
Recency effect, 119
  delay and, 118
Recoding, 10, 31, 184
Recognition, 10
Redundant stimuli, 73-74
Rehearsal, 9, 10, 15, 16, 26, 27, 35, 63, 84, 92, 116-120, 173, 184
  improvement in, 50, 120
  retardate deficit in, 116, 118-120
  spontaneous, 50
  transfer of information by, 115-117
Rehearsal strategy, 16, 29, 31, 34, 68, 112, 113, 173, 184
  effect of learning on, 118
  in retardates, 118-119
Rehearsal strategy deficiency hypothesis, 118
Reinforcement, 104-105
  card sorting and, 128
  choice of task and, 128-129
  defined, 137

  developmental hypothesis of, 128-129
  differential, 138
  differential effect on retardates of, 127
  expectations of, 126
  intermittent, 138
  locus of control and, 129
  partial, 59-60, 62, 81
  predicted by social learning theory, 127-128
  reaction time and, 128
  response strength and, 137
  value of, 126-127
Reinforcement contingencies, 14, 21, 81, 138, 139, 142, 168
  effect of differential, 12
  unintentional, 140
Reinforcement history, inadequate, 139-141
Reinforcement schedules, 138
Relevant cues
  attending to, 81, 167, 173
  temporal contiguity to reward of, 80-82
Relevant stimulus dimensions, 16, 37, 65, 80-82, 147, 152-153, 155-160, 166-167
  attending to, 65, 80, 82, 153, 155-158, 167
  effect of initial tendency to attend to, 153, 155-156, 158-160, 164-167
    of pretraining on, 162
    of reward while attending to, 153, 155-156
  probability of attending to, 152-153, 157-158
    of correct choice response while attending to, 153, 157-158
  rate of increase in probability while attending to, 159
  transfer to new response of, 81
Relevant stimuli
  attending to, 64-65, 80-83
  verbal labels for, 65
Reminiscence, 63
Repetition
  effect of frequency on learning of, 92
  method of modified, 92
  role of, 10
Representation of stimuli in memory, 14, 17-18
Response competition, 122
Response integration, 14
Response selection, 7, 146
Response strength, 98, 137
Retention, 32, 50, 109
  related to age, 32
    to retardation, 32, 166
  symmetry of dot patterns and, 49
  verbalization and, 50
Retention interval, 10, 116-117, 119
Retention loss, 12
Retraining program, 141, 142, 186-187
  attending responses in, 143-144
  computer-assisted instructions in, 145
  fading in, 143-144
  individualization of, 144-145
  using operant conditioning, 141-144
  organizational strategies in, 136
  prevention of errors in, 143
Reversal midplateau, 160, 162-163
  effect of overtraining on, 162
  in five-choice discrimination, 163
  relation of irrelevant dimensions to, 160
Reversal shift, 25, 48, 65-66, 164-165, 169-172, 177-179
  attention theory and, 164-165
  by fast and slow learners, 179
  mediation theory and, 50, 172
  overtraining and, 65-66, 178-180
Reward, 7, 14-15, 37-40, 80-83, 108, *see also* Reinforcement
  anticipation of, 11
  effect on discrimination learning of, 38, 80, 180
  management of, 91-92
  suitable, 142
Rigidity, 132-133

S

Scanning of alternatives, covert, 41
Secondary memory, 115, 117-118
  retardate deficit in, 116, 119

Selective attention, 29, 34, 36-37, 70
  to relevant cues, 170-172
  role of classical conditioning in, 174-175
  to stimuli, 13
  strategies of, 31
Sensorimotor tasks
  inhibition deficit hypothesis and, 63
  massed practice on, 63
  possible effects of verbalization on, 63-64
  previous experience and, 64
Sensory analyzer, 171
Sequential building (easy and hard learning), 124-125
Serial learning, 104-105
  relation of locus of control to, 129
Serial order memory, retrieval skills in, 49
Serial position effects, 49, 68, 110-112, 116-118
  effect of speed of learning on, 111
  Hull's theory of, 111
  relation of rehearsal strategies to, 112
  among retardates, 68, 116-118
  stimulus trace explanation of, 111-112, 119
Shaping of behavior, 62, 138
  by successive approximations, 138-139
Shifts
  extradimensional, 163-165, 169-170, 172, 175-178
  half-reversal, 185
  intradimensional, 163-165, 169-170, 172, 175-178
  mediating responses in, 171-174
  nonreversal, 50, 179
  optional, 50
  reversal, 164-165, 177-179
Short-term forgetting, by retardates, 34
Short-term memory, 9, 109, 184, 188
  in arithmetic, 93
  attention and, 115
  comparative study of, 33, 49, 117
  developmental trends in, 49
  differentiation from long term memory of, 109
  in discrimination learning, 25
  Ellis's theory of, 110, 115-116
  theories of, 115
Short-term memory loss, 12
Short-term retention curve, 9, 119-120
Similarity of CS and US, 75
Social learning theory, 126-131
  postulates and theorems of, 130
  reinforcement history in, 127-128
Spence's theory, 100
  incentive motivation in, 100
Stages of intellectual development, 32
Stimulus change, probability of attending to, 25
Stimulus component, 14
Stimulus compound(s), 10, 35, 36
Stimulus context, 12, 123
Stimulus contingencies, attention to, 20
Stimulus dimensions, 14-16
  defined, 151-152
  relevant, 16, 37, 65, 80-82, 147, 152-153, 155-160, 166-167, 171
  transfer along, 16, 169-171
Stimulus elements, minimum set of, 80
Stimulus generalization, 51, 99, 121-122, 171
Stimulus generalization gradient, 102, 106
  relation to intelligence of, 106
    to stimulus dimension of, 106
Stimulus patterning, 14, 38
Stimulus-response association, 23, 36, 51, 98, 99, 121
  disinhibition of incorrect, 51
Stimulus-response bonds, 7
Stimulus sampling theory, 9, 26, 123, 157, 175
  intertrial variability and, 74
Stimulus scanning, 41
Stimulus selection strategies, 166
Stimulus-stimulus conditioning, 8
Stimulus substitution, 8, 20, 121
Stimulus trace, 18, 110, 111
Stimulus trace hypothesis, 110-114
Strategies, 14
  categorization, 16
  choice point, 16, 24, 26
  coding, 27
  mediation, 28
  organizational, 16
  rehearsal, 16, 20, 26, 29, 31, 68, 112, 113, 173, 184

Structure-of-intellect, 8
Successive acquisition and extinction of response, 15
Successive reversals, 23, 178

T

Talking typewriter, 92
Taylor's manifest anxiety scale, 109
Teaching machines, 87, 89-92
  immediate reward in, 90
  operant conditioning principles in, 89-90
  retention and, 91
  use with retardates of, 90-91
Temporary associations, 33, 35
  consolidation of, 33, 35
  effect of CS on, 33
  redundancy in, 33
Tertiary memory, 115-116, *see also* Long term memory
Transfer, 15, 16, 81
  communality of associations in, 13
  along dimensions, 80, 171
  of observing responses, 171, 177
  to long-term storage, 20, 28, 31, 33, 34, 115-117, 120
Transposition, 50, 66

U

US, in absence of CS, 72-73
  systems of, 39

V

Verbal conditioning, 22, 77-78
  free-responding procedure for, 22
  perception of reward contingencies in, 22
  rate of, 22, 77-78
Verbal control of behavior, 14, 78
Verbal learning, 16, 48, 63, 67-69, 84-86
  developmental trends in, 48
Verbal mediation, 49, 50, 68-69
Verbal mediators, 19, 38, 49, 69-70, 85-86
  developmental studies of, 50-51
  effect
    on recall of, 86
    of recognition of, 86
  in learning from models, 38
  specific training in use of, 69
  spontaneous use of, 70, 86
  transfer of, 69
Verbal reinforcement, 77-78
Vicarious trial and error, 41

West Bengal
Library
Calcutta
B. C. Road

An Academic Press Journal

154.4 EST

LEARNING AND MOTIVA

Editor: F. ROBERT BRUSH

Associate Editors:

A. H. Black · P. L. Carlton · J. S. Myer · J. M. Theios
R. C. Bolles · R. M. Church · D. Premack · F. R. Treichler
B. A. Campbell · R. D. Fitzgerald · W. F. Prokasy · A. R. Wagner

142

This journal publishes experimental and theoretical papers addressed to the analysis of basic phenomena and mechanisms of learning and motivation, broadly defined. Studies using infrahuman subjects are favored over those involving human subjects, except when the latter have special significance for understanding basic principles of learning or motivation.

Volume 1, 1970 (quarterly) $25.00

Postage outside U.S. and Canada $1.60

Contents of Volume 1, Number 1:

Russell M. Church, Carol L. Wooten, and *T. James Matthews,* Discriminative Punishment and the Conditioned Emotional Response.

John A. Bull, III, An Interaction between Appetitive Pavlovian CSs and Instrumental Avoidance Responding.

R. J. W. Mansfield and *Howard C. Rachlin,* The Effect of Punishment, Extinction, and Satiation on Response Chains.

J. D. Keehn and *Maxine Walsh,* Bar-Holding with Negative Reinforcement as a Function of Press- and Release-Shock Intervals.

Dennis J. Delprato and *M. Ray Denny,* Passive Avoidance as a Function of the Duration of Nonshock Confinement.

Walter vom Saal and *Herbert M. Jenkins,* Blocking the Development of Stimulus Control.

G. William Farthing and *Eliot Hearst,* Attention in the Pigeon: Testing with Compounds or Elements.

Jonathan Vaughan and *Allan M. Schrier,* Effect of Locus of Reinforcement on Learning-Set Formation by Rhesus Monkeys.

Paul E. Van Hemel and *James S. Myer,* Control of Food-Motivated Instrumental Behavior in Water-Deprived Rats by Prior Water and Saline Drinking.

Delos D. Wickens, Anthony F. Nield, David S. Tuber, and *Carol Wickens,* Classically Conditioned Compound-Element Discrimination as a Function of Length of Training, Amount of Testing and CS–UCS Interval.

John D. Holmes and *I. Gormezano,* Classical Appetitive Conditioning of the Rabbit's Jaw Movement Response under Partial and Continuous Reinforcement Schedules.

Agu Pert and *M. E. Bitterman,* Reward and Learning in the Turtle.

Academic Press
NEW YORK AND LONDON
111 FIFTH AVENUE, NEW YORK 10003
BERKELEY SQUARE HOUSE, LONDON, W1X 6BA

ISBN 0–12–243550–8